

**STIC-Biotech/Ch mLib**

---

**From:** Nguyen, Dave  
**Sent:** Tuesday, February 12, 2002 6:30 PM  
**To:** STIC-Biotech/ChemLib  
**Subject:** 09/664,225 Peptide search request

Please do a peptide search for the peptide sequences of SEQ ID NOS: 66, 69, 152 and 154.

Thanks.  
Dave Nguyen  
Art Unit 1632  
CM1-12B15  
Mail Box 12E12  
703-305-2024

**of Contact:**  
**Sheppard**

**TYPE OF SEARCH:**

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 2/13/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

**VENDOR/COST(where applic.)**

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 10:01:54 : Search time 62.52 Seconds

(without alignments)  
44.453 Million cell updates/sec

Title: US-09-664-225-154

Sequence: 1 KATLQDVLHLEPQNEIPV 19

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP:REMBL\_17:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	105	12	O9QNP7 human papill
2	96	100.0	105	12	O9QNP5 human papill
3	91	94.8	105	12	O9QNP6 human papill
4	76.5	79.7	106	4	O9Y4Y3 O9Y4Y3 homo sapien
5	76.5	79.7	106	12	O10609 human papill
6	59	61.5	107	12	O81965 human papill
7	56	58.3	108	12	O9WHG1 human papill
8	52	54.2	97	12	O82006 human papill
9	48.5	50.5	110	12	O73452 human papill
10	48	50.0	99	12	O90724 human papill
11	47	49.0	99	12	O37386 common chl
12	47	49.0	369	2	O86548 streptomyc
13	45	46.9	100	12	O9IR58 human papill
14	44.5	46.4	718	13	O9W6G9 xenopus lae
15	44	45.8	374	4	O9H688 homo sapien
16	44	45.8	424	10	O9LIMG9 arabidopsis
17	44	45.8	642	4	O9Y2Z1 homo sapien
18	44	45.8	707	4	O94940 homo sapien
19	44	45.8	758	10	O9SIWO taraxacum o

20	44	45.8	1385	3	O9UUY2 neurospora
21	43.5	45.3	462	5	O02134 caenorhabd
22	43	44.8	104	12	O9JH50 human papil
23	43	44.8	206	4	O9HAB0 homo sapien
24	43	44.8	767	5	O9U234 caenorhabd
25	43	44.8	899	5	O9VFS6 drosophila
26	43	44.8	1393	4	O75872 homo sapien
27	43	44.8	1393	4	O9H2M9 homo sapien
28	42	43.8	98	12	O9QLP4 human papil
29	42	43.8	100	12	O99329 human papil
30	42	43.8	159	12	O83179 cauliflowe
31	42	43.8	288	2	O9PN89 campylobact
32	42	43.8	394	2	O9X213 thermotoga
33	42	43.8	497	2	O9S467 bacillus st
34	42	43.8	4436	1	O58659 pyrococcus
35	42	43.8	5038	4	O9NPK1 homo sapien
36	41.5	43.2	1620	11	O9QXC0 mus musculi
37	41.5	43.2	2069	11	O9ESD7 mus musculi
38	41	42.7	132	2	O55927 synchocyst
39	41	42.7	211	2	O9KRP2 synchocyst
40	41	42.7	260	2	P73620 synchocyst
41	41	42.7	264	2	P72822 synchocyst
42	41	42.7	268	2	P73838 synchocyst
43	41	42.7	295	5	O9N3E1 caenorhabd
44	41	42.7	345	4	O9UJ22 homo sapien
45	41	42.7	345	4	O9NRA1 homo sapien

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	105 AA.
O9QNP7				
AC O9QNP7				
DT 01-MAY-2000 (TREMBLrel. 13, Created)				
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)				
DE E7 PROTEIN.				
GN E7.				
OS Human papillomavirus type 18.				
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;				
OC Papillomavirus.				
OX NCBI_TaxID=10582;				
RN [1]				
RP SEQUENCE FROM N.A.				
RA Laasari M., Gul'ko L., Vinokurova S., Kissel'ova N., Veiko V.,				
RA Kissel'ev F.;				
RT "Cloning of E6 and E7 Genes of Human Papilloma Virus Type 18 and				
RT Transformation Potential of E7 Gene and its Mutants.";				
RL Virus Genes 182:139-149(1999).				
DR EMBL:Y18491; CAB53097.1.-.				
DR Interpro: IPR000148; Papyl_E7.				
DR Pfam: PF00527; E7: 1.				
SQ SEQUENCE 105 AA; 12010 MW; 24799BB534D4496A CRC64;				
Query Match	100.0%;	Score 96;	DB 12;	Length 105;
Best Local Similarity	100.0%;	Pred. No. 3.3e-08;		
Matches 19;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
OY 1 KATLQDVLHLEPQNEIPV 19				
DB 5 KATLQDVLHLEPQNEIPV 23				
RESULT 2				
O9QNP5				
ID O9QNP5		PRELIMINARY:	PRT:	105 AA.
AC O9QNP5				
DT 01-MAY-2000 (TREMBLrel. 13, Created)				
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)				

DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 18.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10582;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laessli M., Gul'ko L., Vinokurova S., Kisseljova N., Velko V.,  
 Kisseljov F.;  
 RT "Cloning of E6 and E7 genes of Human Papilloma Virus Type 18 and  
 RT Transformation Potential of E7 Gene and Its Mutants.";  
 RL Virus Genes 182:139-149(1999).  
 DR EMBL; Y18493; CAB53099.1; -;  
 DR InterPro; IPR000148; PapvL\_E7.  
 DR Pfam; PF00527; E7; 1.  
 SQ SEQUENCE 105 AA; 11952 MW; 247BF4486BBB4FA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-08;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLQDIVLHPONEIPV 19  
 :|||||:|||||:|||||:  
 Db 5 KATLQDIVLHPONEIPV 23

RESULT 3  
 ID Q9ONP6 PRELIMINARY; PRT; 105 AA.  
 AC Q9ONP6;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 18.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10582;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laessli M., Gul'ko L., Vinokurova S., Kisseljova N., Velko V.,  
 Kisseljov F.;  
 RT "Cloning of E6 and E7 genes of Human Papilloma Virus Type 18 and  
 RT Transformation Potential of E7 Gene and Its Mutants.";  
 RL Virus Genes 182:139-149(1999).  
 DR EMBL; Y18492; CAB53098.1; -;  
 DR InterPro; IPR000148; PapvL\_E7.  
 DR Pfam; PF00527; E7; 1.  
 SQ SEQUENCE 105 AA; 12009 MW; C497955DDAD4A960 CRC64;

Query Match 94.8%; Score 91; DB 12; Length 105;  
 Best Local Similarity 94.7%; Pred. No. 2.1e-07;  
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLQDIVLHPONEIPV 19  
 :|||||:|||||:|||||:  
 Db 5 KATLQDIVLHPONEIPV 23

RESULT 4  
 ID Q9V4Y3 PRELIMINARY; PRT; 106 AA.  
 AC Q9V4Y3;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE E7 PROTEIN.  
 GN HPV45 E7.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sastre-Garau X., Favre M., Couturier J., Orth G.;  
 RT "Distinct patterns of alterations of myc genes associated with  
 RT integration of HPV16 or HPV45 dna in two genital tumors.";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; A1242956; CAB44707.1; -;  
 DR InterPro; IPR000148; PapvL\_E7.  
 DR Pfam; PF00527; E7; 1.  
 SQ SEQUENCE 106 AA; 11977 MW; 1912D1DE742F6BED CRC64;

Query Match 79.7%; Score 76.5; DB 4; Length 106;  
 Best Local Similarity 80.0%; Pred. No. 5e-05;  
 Matches 16; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 KATLQDIVLHPONEI-PV 19  
 :|||||:|||||:|||||:  
 Db 5 RATLQDIVLHPONEIDPV 24

RESULT 5  
 ID O10609 PRELIMINARY; PRT; 106 AA.  
 AC O10609;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ONCOPROTEIN E7.  
 GN E7.  
 OS Human papillomavirus type 45.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10593;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-1C4 HPV45 VARIANT;  
 RC Sastre-Garau X., Favre M., Couturier J., Orth G.;  
 RL J. Gen. Virol. 0:0-0(0);  
 DR EMBL; Y13218; CA73661.1; -;  
 DR InterPro; IPR000148; PapvL\_E7.  
 DR Pfam; PF00527; E7; 1.  
 SQ SEQUENCE 106 AA; 11977 MW; 1912D1DE742F6BED CRC64;

Query Match 79.7%; Score 76.5; DB 12; Length 106;  
 Best Local Similarity 80.0%; Pred. No. 5e-05;  
 Matches 16; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 KATLQDIVLHPONEI-PV 19  
 :|||||:|||||:|||||:  
 Db 5 RATLQDIVLHPONEIDPV 24

RESULT 6  
 ID Q81965 PRELIMINARY; PRT; 107 AA.  
 AC Q81965;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ORF POUAFIVE E7.  
 OS Human papillomavirus type 5.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10578;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94303229; PubMed=8030272;  
 RA Rho J., Roy-Burman A., Kim H., De Villiers E.M., Matsukura T.,



```

RA Choe J.;
RT "Nucleotide sequence and phylogenetic classification of human
RT Papillomavirus type 59."
RL Virology 203:158-161(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Choe J.;
RL Submitted (FEB-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL: X77858; CA54850.1; -.
DR InterPro: IPR000148; Papvi_E7.
DR Pfam: PF00527; E7; 1.
SQ SEQUENCE 107 AA; 12042 MW; A8F2B6973D727163 CRC64;

Query Match
Best Local Similarity 61.5%; Score 59; DB 12; Length 107;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KATLQDIVLHLEPQNEI 15
DB 5 KATLQDIVLHLEPQNEI 19

RESULT 7
Q9WHG1 PRELIMINARY; PRT; 108 AA.
AC 09WHG1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE TRANSFORMING PROTEIN E7.
GN E7.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HLT7474-S;
RA "Chow V.T.K., Leong W.F.;
RT "Complete nucleotide sequence of a novel genital human papillomavirus
RT HLT7474-S from Singapore."
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF131950; AAD24182.1; -.
DR InterPro: IPR000148; Papvi_E7.
DR Pfam: PF00527; E7; 1.
SQ SEQUENCE 108 AA; 12355 MW; B1F84D7DA0F384FB CRC64;

Query Match
Best Local Similarity 58.3%; Score 56; DB 12; Length 108;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATLQDIVLHLEPQNEI 17
DB 5 KPTVHDIVLDEPYNEV 21

RESULT 8
Q82006 PRELIMINARY; PRT; 97 AA.
AC Q82006;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE E6, E7, E1, E2, E4, L2, AND L1 GENES.
GN E7.
OS Human papillomavirus type 7.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10620;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=96213783; PubMed=8635859;
RA Voelter C., He Y., Dellus H., Roy-Burman A., Greenspan J.S.,
RA Greenspan D., De Villiers E.M.;
RT "Novel HPV types present in oral papillomatous lesions from patients
RT with HIV infection."
RL Int. J. Cancer 66:453-456(1996).
DR EMBL: X94165; CAA63883.1; -.
DR InterPro: IPR000148; Papvi_E7.
DR Pfam: PF00527; E7; 1.
SQ SEQUENCE 97 AA; 10970 MW; 651D0345D048F022 CRC64;

Query Match
Best Local Similarity 54.2%; Score 52; DB 12; Length 97;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KATLQDIVLHLEPQNEI 17
DB 5 KTLQDIVLHLEPQNEI 21

RESULT 9
O73452 PRELIMINARY; PRT; 110 AA.
AC 073452;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 68.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=45240;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98090464; PubMed=9427755;
RA Reuter S., Bartelmann M., Vogt M., Gelsen C., Napjerski I., Kahn T.,
RA Dellus H., Lichter P., Wetz S., Korn B., Schwarz E.;
RT "APM-1, a novel human gene, identified by aberrant co-transcription
RT with papillomavirus oncogenes in a cervical carcinoma cell line,
RT encodes a BTB/POZ-zinc finger protein with growth inhibitory
RT activity."
RL EMBO J. 17:215-222(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374616; PubMed=1716694;
RA Reuter S., Dellus H., Kahn T., Hofmann B., Zur Hausen H., Schwarz E.;
RT "Characterization of a novel human papillomavirus DNA in the cervical
RT carcinoma cell line ME180."
RL J. Virol. 65:5564-5568(1991).
DR EMBL: Y14591; CAA74932.1; -.
DR InterPro: IPR000148; Papvi_E7.
DR Pfam: PF00527; E7; 1.
SQ SEQUENCE 110 AA; 12700 MW; 90ECC0E01476BDB CRC64;

Query Match
Best Local Similarity 50.5%; Score 48.5; DB 12; Length 110;
Matches 13; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 KATLQDIVLHLEPQNEI 19
DB 5 KPTVQDIVLHLEPQNEI 24

RESULT 10
O90724 PRELIMINARY; PRT; 99 AA.
AC 090724;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

```

DE ORF E7.  
 OS Human papillomavirus type 67.  
 OC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 OC Papillomavirus.  
 CX NCBI\_Taxid=37120;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kiril Y., Iwamoto S., Matsukura T.;  
 RT "Complete DNA sequence of human Papillomavirus type 67 isolated from a  
 RT vaginal intraepithelial neoplasia."  
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D21208; BAA28853.1; -  
 DR InterPro: IPR00148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 SO SEQUENCE 99 AA; 11156 MW; 81A4A636D7588C44 CRC64;

Query Match 50.0%; Score 48; DB 12; Length 99;  
 Best Local Similarity 64.3%; Pred. No. 2.1;  
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KATLDIVLHLEPQ 14  
 ||||| :||| :  
 Db 5 KATLDIVLHLEPQ 18

RESULT 11  
 ID 037386 PRELIMINARY; PRT; 99 AA.  
 AC 037386;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
 DE COMPLETE GENOME.  
 GN E7.  
 OS Common chimpanzee papillomavirus 1.  
 OC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 OC Papillomavirus.  
 CX NCBI\_Taxid=6261;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RA Scinicariello F., Soza I., Brasky K.M., Hilliard J.K.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF020905; AAB7105.1; -  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 SO SEQUENCE 99 AA; 10858 MW; 738500347E87AFE CRC64;

Query Match 49.0%; Score 47; DB 12; Length 99;  
 Best Local Similarity 81.8%; Pred. No. 3.1;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 TLQDIVLHLEP 13  
 ||||| :||| :  
 Db 7 TLQDIVLHLEP 17

RESULT 12  
 ID 086548 PRELIMINARY; PRT; 369 AA.  
 AC 086548;  
 DT 01-NOV-1998 (TEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
 DE PUTATIVE MEMBRANE PROTEIN.  
 GN SCIF2.11.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 CX NCBI\_Taxid=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Parkhill J., Barrell B.G., Randalream M.A.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redegbachi M., Kleser H.M., Denapate D., Eichner A., Cullum J.,  
 RA Khashil H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
 RL MOL. MICROBIOL. 21:77-96(1996).  
 DR EMBL; AL031350; CAA20502.1; -  
 SO SEQUENCE 369 AA; 39663 MW; 8C50A4C2922D6E13 CRC64;

Query Match 49.0%; Score 47; DB 2; Length 369;  
 Best Local Similarity 61.1%; Pred. No. 12;  
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 ATLQDIVLHLEPQNEIPV 19  
 ||||| :||| :  
 Db 246 ATLQDIVLHLEPQNEIPV 263

RESULT 13  
 ID 091R58 PRELIMINARY; PRT; 100 AA.  
 AC 091R58;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
 DE DNA, COMPLETE GENOME.  
 GN E7.  
 OS Human papillomavirus type 82.  
 OC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 OC Papillomavirus.  
 CX NCBI\_Taxid=129724;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RA Sata T., Matsukura T.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20087389; PubMed=10618284;  
 RA Kino N., Sata T., Sato Y., Sugase M., Matsukura T.;  
 RT "Molecular cloning and nucleotide sequence analysis of a novel human  
 RT papillomavirus (type 82) associated with vaginal intraepithelial  
 RT neoplasia."  
 RL Clin. Diagn. Lab. Immunol. 7:91-95(2000).  
 DR EMBL; AB027021; BAA90736.1; -  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 SO SEQUENCE 100 AA; 11325 MW; 47E8C0D0FD67A81D CRC64;

Query Match 46.9%; Score 45; DB 12; Length 100;  
 Best Local Similarity 71.4%; Pred. No. 6.6;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 LQDIVLHLEPQNEI 17  
 ||||| :||| :  
 Db 8 LQDIVLHLEPQNEI 21

RESULT 14  
 ID 09W6G9 PRELIMINARY; PRT; 718 AA.

AC Q9W6G9;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)  
 DE SPINDLE CHECKPOINT PROTEIN XMAD1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99003289; PubMed=9786942;  
 RA Chen R.-H., Shevchenko A., Mann M., Murray A.W.;  
 RT "Spindle checkpoint protein Xmad1 recruits Xmad2 to unattached  
 RT kinetochores.";  
 RL J. Cell Biol. 143:283-295(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Chen R.-H., Shevchenko A., Mann M., Murray A.W.;  
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF126481; AAD25081.1; -; 044FBCAD1459A51A CRC64;  
 SQ SEQUENCE 718 AA; 83210 MW; 044FBCAD1459A51A CRC64;

Query Match 46.4%; Score 44.5; DB 13; Length 718;  
 Best Local Similarity 58.8%; Pred. No. 62;  
 Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 3 TLQDIV-LHLEPONEIP 18  
 ||:|:|:|:|:|  
 Db 685 TLRDFDLHLHHQNSIP 701

RESULT 15  
 Q9H688 PRELIMINARY; PRT; 374 AA.  
 AC Q9H688;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE CDNA: FLJ22490 FIS, CLONE HRC10983.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,  
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
 RT "NEO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AK026143; BAB15375.1; -;  
 SQ SEQUENCE 374 AA; 43215 MW; BB59352711CEDA86 CRC64;

Query Match 45.8%; Score 44; DB 4; Length 374;  
 Best Local Similarity 43.8%; Pred. No. 38;  
 Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 4 LODIVLLEPONEIPV 19  
 ||:|:|:|:|:|  
 Db 80 LODRLHMDSDDEIPI 95

Search completed: February 13, 2002, 10:01:54  
 Job time: 253 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 09:59:18 ; Search time 66.5 seconds

(without alignments)  
46.783 Million cell updates/sec

Title: US-09-664-225-152

Perfect score: 223  
Sequence: 1 RRPYKLPDLCTELNTSLQDI.....TCVYCKTVLETFEVEFFAFK 42

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.Geneseq-1101:\*

1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	223	100.0	42	22	AA895963
2	223	100.0	119	22	AA896051
3	223	100.0	158	16	AA879656
4	223	100.0	158	16	AA863866
5	223	100.0	158	21	AA803176
6	223	100.0	158	21	AA803176
7	223	100.0	158	22	AA898427
8	223	100.0	236	22	AA896052
9	223	100.0	271	13	AA827728
10	223	100.0	278	20	AA825385
11	223	100.0	278	20	AA825385
					Prot.DI/3-E6-His/H

12	223	100.0	383	20	AA825386	HPV fusion protein
13	223	100.0	383	20	AA825386	Prot.DI/3-E6-E7-Hi
14	187	83.9	158	22	AA898440	Human papillomavirus
15	160	71.7	32	12	AA814749	Seroreactive epitope
16	130	58.3	151	21	AA857808	HPV-16 E6 protein
17	130	58.3	158	13	AA822766	HPV E6 peptide, S
18	130	58.3	158	21	AA82462	Human papillomavirus
19	130	58.3	158	22	AA898420	Human papillomavirus
20	130	58.3	162	18	AA835741	Human papillomavirus
21	130	58.3	162	18	AA835742	Human papillomavirus
22	130	58.3	172	17	AA897563	Human papillomavirus
23	130	58.3	188	16	AA863865	HPV16 E6/E7 protei
24	130	58.3	243	20	AA899369	Papillomavirus E6/
25	130	58.3	263	13	AA827725	HPV 16 E6 protein
26	130	58.3	266	17	AA897561	Human papilloma vi
27	130	58.3	273	20	AA825376	HPV fusion protein
28	130	58.3	273	20	AA825379	Prot.DI/3-E6-His/H
29	130	58.3	292	20	AA825379	HPV fusion protein
30	130	58.3	292	20	AA825379	CLYTA-E6-His prote
31	130	58.3	371	20	AA825377	HPV fusion protein
32	130	58.3	371	20	AA825377	Prot.DI/3-E6-E7-Hi
33	130	58.3	390	20	AA825381	HPV fusion protein
34	130	58.3	390	20	AA825381	CLYTA-E6E7-His pro
35	130	58.3	395	14	AA840919	HPV E6 region prod
36	105	47.1	149	22	AA898447	Human papillomavir
37	104	46.6	30	22	AA830942	Peptide comprising
38	104	46.6	30	22	AA831016	Polypeptidic fragm
39	104	46.6	30	22	AA831108	A polypeptidic fra
40	91	40.8	149	22	AA898434	Human papillomavir
41	88	39.5	150	22	AA898395	Human papillomavir
42	88	39.5	150	22	AA898404	Human papillomavir
43	88	39.5	368	18	AA801808	Human papillomavir
44	88	39.5	375	18	AA801809	Human papillomavir
45	88	39.5	465	18	AA801810	Human papillomavir

## ALIGNMENTS

RESULT 1	
AA895963	standard; peptide; 42 AA.
ID	AA895963
AC	AA895963
XX	
DT	25-JUN-2001 (first entry)
XX	
DE	HPV 18 E6 protein ftragment SEQ ID 152.
XX	
XX	
KW	Epitope: tumour antigen; antiviral; immunostimulatory; cervical cancer;
KW	human papillomavirus-associated disease; condyloma; cervical dysplasia;
KW	cervical dysplasia; major histocompatibility complex; MHC I.
XX	
OS	Human papillomavirus.
PN	WO200119408-A1.
XX	
PD	22-MAR-2001.
XX	
PE	18-SEP-2000; 2000WO-US25559.
XX	
PR	16-SEP-1999; 99US-0154665.
PR	16-SEP-1999; 99US-0398534.
PR	09-DEC-1999; 99US-0169846.
XX	
PR	09-DEC-1999; 99US-0458173.
XX	
PA	(ZYCO-) ZYCOS INC.
XX	
PI	Hedley ML, Urban RC, Chicz RM;
XX	
DR	WPI; 2001-265996/27.
XX	
PT	Novel nucleic acids encoding polypeptide polypeptides containing

PT multiple epitopes from one or more proteins, useful for treating tumors  
PT and as vaccines against pathogenic agents -  
PS Claim 46; Page 10; 64pp; English.

CC This invention relates to polynucleotides encoding a hybrid polypeptide  
CC comprising a signal sequence and three segments that are either  
CC contiguous or separated by a spacer amino acid or spacer peptide. The  
CC invention specifically details polynucleotides encoding a polypeptide  
CC peptide where the peptide segments are tumor antigens or a naturally  
CC occurring protein of a pathogenic agent. The polypeptide peptides exhibit  
CC antiviral and immunostimulatory activity. The polynucleotide and  
CC polypeptide peptides are useful for eliciting an immune response in a  
CC mammal. The polynucleotide and protein are useful as vaccines for  
CC treating tumors and pathogenic infections. The polynucleotide is also  
CC useful for preventing or treating human papillomavirus (HPV)-associated  
CC diseases, particularly exophytic condyloma, flat condyloma, cervical  
CC cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV  
CC infection, cervical dysplasia, high grade squamous intraepithelial  
CC lesions, and anal HPV infection. The polynucleotide and polypeptide are  
CC useful for generating or enhancing prophylactic or therapeutic immune  
CC response against pathogens, tumors or autoimmune diseases in a  
CC population of individuals having diverse MHC allotypes, as positive  
CC controls in T cell stimulation assays in vitro, and as tools to  
CC understand processing of epitopes within cells. Peptides  
CC AAB95894 - AAB96037 and AAB96044 - AAB96048 represent major  
CC histocompatibility complex I (MHC I) associated tumor and pathogen  
CC antigens. The peptides can be used as part of the polypeptide proteins of  
CC the invention. Also included are examples of the polypeptide proteins  
CC represented by AAB96050 - AAB96052, and localisation signal peptides  
CC AAB96038 - AAB96043 and AAB96049 which can be used in the construction of  
CC the polypeptide peptides.

CC Sequence 42 AA:

Query Match 100.0%; Score 223; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1e-23;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DY 1 RRPYKLPDCLTELNSTLSDIFETCYCKTVLETFEFAFK 42  
DB 1 RRPYKLPDCLTELNSTLSDIFETCYCKTVLETFEFAFK 42

#### RESULT 2

AAB96051  
ID AAB96051 standard; Peptide: 119 AA.

AC AAB96051;

DT 25-JUN-2001 (first entry)

DE polypeptide polypeptide SEQ ID 159.

KW Epitope: tumour antigen; antiviral; immunostimulatory; cervical cancer;  
KW human papillomavirus-associated disease; condyloma; cervical dysplasia;  
KW cervical dysplasia; major histocompatibility complex; MHC I.

OS Synthetic.

PN WO200119408-A1.

PD 22-MAR-2001.

PE 18-SEP-2000; 2000WO-US255559.

PR 16-SEP-1999; 99US-0154665.

PR 16-SEP-1999; 99US-0398534.

PR 09-DEC-1999; 99US-0169846.

PR 09-DEC-1999; 99US-0458173.

PR (ZTCO-) ZTCOS INC.

XX Hedley ML, Urban RC, Chicz RM;  
PI WPL; 2001-265996/27.

PT Novel nucleic acids encoding polypeptide polypeptides containing  
PT multiple epitopes from one or more proteins, useful for treating tumors  
PT and as vaccines against pathogenic agents -  
PS Disclosure; Fig 4; 64pp; English.

CC This invention relates to polynucleotides encoding a hybrid polypeptide  
CC comprising a signal sequence and three segments that are either  
CC contiguous or separated by a spacer amino acid or spacer peptide. The  
CC invention specifically details polynucleotides encoding a polypeptide  
CC peptide where the peptide segments are tumor antigens or a naturally  
CC occurring protein of a pathogenic agent. The polypeptide peptides exhibit  
CC antiviral and immunostimulatory activity. The polynucleotide and  
CC polypeptide peptides are useful for eliciting an immune response in a  
CC mammal. The polynucleotide and protein are useful as vaccines for  
CC treating tumors and pathogenic infections. The polynucleotide is also  
CC useful for preventing or treating human papillomavirus (HPV)-associated  
CC diseases, particularly exophytic condyloma, flat condyloma, cervical  
CC cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV  
CC infection, cervical dysplasia, high grade squamous intraepithelial  
CC lesions, and anal HPV infection. The polynucleotide and polypeptide are  
CC useful for generating or enhancing prophylactic or therapeutic immune  
CC response against pathogens, tumors or autoimmune diseases in a  
CC population of individuals having diverse MHC allotypes, as positive  
CC controls in T cell stimulation assays in vitro, and as tools to  
CC understand processing of epitopes within cells. Peptides  
CC AAB95894 - AAB96037 and AAB96044 - AAB96048 represent major  
CC histocompatibility complex I (MHC I) associated tumor and pathogen  
CC antigens. The peptides can be used as part of the polypeptide proteins of  
CC the invention. Also included are examples of the polypeptide proteins  
CC represented by AAB96050 - AAB96052, and localisation signal peptides  
CC AAB96038 - AAB96043 and AAB96049 which can be used in the construction of  
CC the polypeptide peptides.

CC Sequence 119 AA:

Query Match 100.0%; Score 223; DB 22; Length 119;  
Best Local Similarity 100.0%; Pred. No. 3.2e-23;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 1 RRPYKLPDCLTELNSTLSDIFETCYCKTVLETFEFAFK 42

DB 1 RRPYKLPDCLTELNSTLSDIFETCYCKTVLETFEFAFK 42

#### RESULT 3

AAR79656  
ID AAR79656 standard; Protein: 158 AA.

AC AAR79656;

DT 06-DEC-1995 (first entry)

DE HPV-18 E6 protein.

KW Ubiquitin-conjugating enzyme; HPV-18 E6 protein; cell cycle;  
KW cell proliferation; cancer; psoriasis; fibrosis.

OS Homo sapiens.

PN WO9518974-A.

PD 13-JUL-1995.

PE 04-JAN-1995; 95WO-US00164.

PR 13-SEP-1994; 94US-0305520.

PR 04-JAN-1994: 94US-0176937.  
 PR 23-MAY-1994: 94US-0247904.  
 PR 27-MAY-1994: 94US-0250795.  
 XX  
 PA (MITO-) MITOTIX INC.  
 XX  
 PI Cottarel G, Draetta G, Eckstein JW, Gyuris J, Rolfe M;  
 XX  
 DR WPI: 1995-255137/33.  
 DR N-PSDB: AAO97848.  
 XX  
 PT Identifying inhibitors of ubiquitin mediated proteolysis of cell cycle  
 XX regulatory proteins - also new ubiquitin conjugating enzymes, their  
 PT related nucleic acid, vectors, antibodies etc., useful for regulating  
 PT e.g. cell proliferation  
 XX  
 PS Disclosure: Page 100-101; 157pp; English.  
 XX  
 CC HPV-18 E6 cDNA (given in AAO97848) was amplified from a HeLa cell  
 CC cDNA library using the primers given in AAO97846-47. The gene  
 CC was subcloned into a baculovirus vector for expression of  
 CC recombinant E6 in Sf9 insect cells for use as a component of an  
 CC in vitro ubiquitin conjugating system.  
 XX  
 SQ Sequence 158 AA;

Query Match 100.0%; Score 223; DB 16; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-23;  
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 RRPYKLPDCTELNTSLQDIETTCVYCKTVLETFEVEFAFK 42  
 Db 9 rrpYkLpDcTeLnTsLqdiEtTcVcKtVleTfeVeFaFk 50

RESULT 4  
 AAR63866  
 ID AAR63866 standard; Protein; 158 AA.  
 XX  
 AC AAR63866;  
 XX  
 DT 28-JUN-1995 (first entry)  
 XX  
 DE HPV18 E6/E7 proteins.  
 XX  
 KM HPV, HPV18; E6 protein; E7 protein; diagnosis; cervical dysplasia;  
 KM cervix cancer.  
 XX  
 OS Human papillomavirus-strain 18.  
 XX  
 PN WO9426934-A.  
 XX  
 PD 24-NOV-1994.  
 XX  
 PF 06-MAY-1994: 94MO-US05085.  
 XX  
 PR 06-MAY-1993: 93US-0058920.  
 XX  
 PA (BAXT ) BAXTER DIAGNOSTICS INC.  
 XX  
 PI Brown JT;  
 XX  
 DR WPI: 1995-006821/01.  
 DR P-PSDB: AAO75471.  
 XX  
 PT Human papilloma virus detection assay - by amplification using  
 PT self sustained sequence replication and hybridisation with a  
 PT detector probe  
 XX  
 PS Disclosure: Page 27-28; 79pp; English.  
 XX  
 CC The sequences of the E6 and E7 polypeptide-encoding regions of human

CC Papillomavirus (HPV) 16 and 18 are given in AAO75470-71 and the  
 CC encoded proteins in AAR63865-66, respectively. Probes and primers  
 CC based on these sequences were used for HPV infection diagnosis;  
 CC expression of E6 and E7 is diagnostic for cervical cancer or pre-  
 CC malignant states.  
 XX  
 SQ Sequence 158 AA;

Query Match 100.0%; Score 223; DB 16; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-23;  
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 RRPYKLPDCTELNTSLQDIETTCVYCKTVLETFEVEFAFK 42  
 Db 9 rrpYkLpDcTeLnTsLqdiEtTcVcKtVleTfeVeFaFk 50

RESULT 5  
 AAY39968  
 ID AAY39968 standard; Protein; 158 AA.  
 XX  
 AC AAY39966;  
 XX  
 DT 15-DEC-1999 (first entry)  
 XX  
 DE HPV-18 E6 protein sequence.  
 XX  
 KM ubiquitin conjugating enzyme; UBCE; ubiquitin-mediated proteolysis;  
 KM cell-cycle regulatory protein; ubiquitination inhibitor; atherosclerosis;  
 KM proliferative disorder; cancer; restenosis; tissue connective disorder;  
 KM wound healing; fibrosis disorder; rheumatoid arthritis; scleroderma;  
 KM insulin dependent diabetes mellitus; glomerulonephritis; cirrhosis;  
 KM diagnosis; therapy; E6.  
 XX  
 OS Human papillomavirus.  
 XX  
 PN US5968761-A.  
 XX  
 PD 19-OCT-1999.  
 XX  
 PF 07-JUN-1995: 95US-0486663.  
 XX  
 PR 04-JAN-1994: 94US-0176937.  
 PR 23-MAY-1994: 94US-0247904.  
 PR 27-MAY-1994: 94US-0250795.  
 PR 13-SEP-1994: 94US-0305520.  
 XX  
 PA (MITO-) MITOTIX INC.  
 XX  
 PI Chiu MI, Cottarel G, Berlin V, Damagnez V, Draetta G, Rolfe M;  
 XX  
 DR WPI: 1999-590402/50.  
 DR N-PSDB: AAZ27568.  
 XX  
 PT Identifying ubiquitination inhibitors using novel ubiquitin conjugating  
 PT enzymes -  
 XX  
 PS Example 2; Column 89-92; 61pp; English.  
 XX  
 CC This sequence is the human papillomavirus E6 protein. The invention  
 CC relates to assays for identifying an inhibitor of ubiquitin-mediated  
 CC proteolysis of a cell-cycle regulatory protein comprising contacting a  
 CC candidate agent with an ubiquitin-conjugating system and measuring the  
 CC level of ubiquitination. The ubiquitin-conjugating system comprises:  
 CC (a) a reconstituted protein mixture including a ubiquitin conjugating  
 CC enzyme (UBCE) produced by the expression of a nucleic acid which  
 CC hybridizes under high stringency conditions to human UBCE. *Candida*  
 CC albicans UBCE, or *Schizosaccharomyces pombe* UBCE coding sequences;  
 CC (b) a regulatory protein; and (c) ubiquitin. The polynucleotides are  
 CC useful for identifying ubiquitination inhibitors. The polynucleotides, be  
 CC polypeptides, antisense compounds and antibodies against them may also  
 CC be useful for the treatment and/or diagnosis of proliferative disorders

CC (e.g. cancer, atherosclerosis, or restenosis), tissue connective  
 CC disorders, controlling wound healing, and disorders characterized by  
 CC fibrosis (e.g. rheumatoid arthritis, insulin dependent diabetes mellitus,  
 CC glomerulonephritis, cirrhosis, and scleroderma).  
 XX  
 SO Sequence 158 AA;

Query Match 100.0%; Score 223; DB 20; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-23;  
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 RRPYKLPDLCTELNTSLQDIEITCVCKTVLELTFEVEFAFK 42  
 Db 9 rrpYkPlpdlcElntslqdlEtCvckTvlElTfEfEfAfK 50

RESULT 6  
 ID AAB03176 standard; Protein: 158 AA.  
 XX AAB03176;  
 AC AAB03176;  
 XX 23-OCT-2000 (first entry)  
 DT  
 XX HPV-18 E6 protein.  
 DE  
 XX HPV-18 E6; ubiquitin mediated proteolysis;  
 KW human; cellular protein half life; ubiquitination inhibitor; p53;  
 KW cyclin; cell cycle regulator; myc deregulation; human papillomavirus;  
 KW HPV-18 E6 protein; cervical cancer; skin cancer; epidermal hyperplasia;  
 KW epidermal neoplasia; psoriasis; connective tissue disorder;  
 KW wound healing; cytostatic; antiproliferative; anticancer; antipsoriatic.  
 KW  
 XX Human papillomavirus 18.  
 OS  
 XX US6068982-A.  
 PN  
 XX 30-MAY-2000.  
 PD  
 XX 17-DEC-1996; 96US-0767942.  
 PF  
 XX 07-JUN-1995; 95US-0486663.  
 PR 04-JAN-1994; 94US-0176937.  
 PR 23-MAY-1994; 94US-0247904.  
 PR 27-MAY-1994; 94US-0250795.  
 PR 13-SEP-1994; 94US-0305520.  
 XX  
 PA (MITO-) MITOTIX INC.  
 XX  
 PI Chiu MI, Cottarel G, Berlin V, Draetta G, Damagnez V, Rolfe M;  
 DR WPI; 2000-410854/35.  
 DR N-PSDB: AAA61623.  
 XX  
 PT Identifying an inhibitor of ubiquitin mediated proteolysis of  
 PT regulatory protein for treating cancers involves measuring  
 PT ubiquitination levels of the protein in the presence of candidate agent  
 PT in an eukaryotic cell  
 PT  
 XX  
 XX Example 2; Column 97-100; 73pp; English.

The invention relates to a method of identifying an inhibitor of  
 CC ubiquitin-mediated proteolysis of a cell cycle regulatory protein  
 CC comprising contacting an engineered eukaryotic cell with a candidate  
 CC agent. The eukaryotic cells is engineered to express a recombinant  
 CC human, *Candida albicans* or *Schizosaccharomyces pombe* ubiquitin-  
 CC conjugating enzyme (AAB03169-B03171), a cell cycle regulatory protein  
 CC (such as p53) and ubiquitin. The specification also discloses novel  
 CC *Candida albicans* and *Schizosaccharomyces pombe* ubiquitin-conjugating  
 CC enzymes, caddC (AAB03170, AAB03171), and two novel human  
 CC ubiquitin-conjugating enzymes, hUbC1 and rapUBC (AAB03169, AAB03173).  
 CC The ubiquitin-mediated proteolysis system is the major pathway for the

CC selective, controlled degradation of intracellular proteins in  
 CC eukaryotic cells. In particular, this system controls the half-lives of  
 CC cellular proteins, and is important in controlling the levels of proteins  
 CC involved in cell cycle progression. Alterations in the ubiquitination of  
 CC these proteins may therefore play a role in the development of cancers.  
 CC For example, human papillomaviruses such as HPV-18 encode a transforming  
 CC protein, E6 (AAB03176), which combines with a cellular E6-associated  
 CC protein (E6-AP; AAB03177) to stimulate the ubiquitination of p53, thus  
 CC targeting it for degradation. The ubiquitination inhibitors identified  
 CC according to the method of the invention are useful for treatment of  
 CC cervical cancers and connective tissue disorders and for controlling the  
 CC wound healing process. They are also useful in treatment of hyperplastic  
 CC epidermal conditions such as psoriasis, neoplastic epidermal conditions,  
 CC skin cancers e.g., basal cell carcinomas, squamous cell carcinomas. The  
 CC inhibitors are useful for deregulating myc expression and rendering the  
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of  
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-  
 CC mediated degradation of cyclins are useful as antiproliferative agents.  
 CC The present sequence represents HPV-18 E6 protein.  
 CC  
 XX  
 SO Sequence 158 AA;

Query Match 100.0%; Score 223; DB 21; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-23;  
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 RRPYKLPDLCTELNTSLQDIEITCVCKTVLELTFEVEFAFK 42  
 Db 9 rrpYkPlpdlcElntslqdlEtCvckTvlElTfEfEfAfK 50

RESULT 7  
 ID AAB98427 standard; Protein: 158 AA.  
 XX AAB98427;  
 AC AAB98427;  
 XX 22-AUG-2001 (first entry)  
 DT  
 XX Human papillomavirus protein HPV18 E6.  
 DE  
 XX Human papillomavirus; human leukocyte antigen; HLA; immune response;  
 KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;  
 KW neoplastic growth; antiviral.  
 KW  
 XX Human papillomavirus.  
 OS  
 XX WO200141799-A1.  
 PN  
 XX 14-JUN-2001.  
 PD  
 XX 11-DEC-2000; 2000WO-US33549.  
 PF  
 XX 10-DEC-1999; 99US-0172705.  
 PR 15-AUG-2000; 2000US-0641528.  
 PR  
 XX (EPTM-) EPTMKONE INC.  
 PA  
 XX Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;  
 PI WPI; 2001-381497/40.  
 DR  
 XX  
 XX An isolated human papilloma virus (HPV) epitope, useful in vaccines for  
 PT treating HPV infections -  
 PT  
 XX  
 PS Disclosure; Page 22; 756pp; English.  
 XX  
 CC The present invention describes an isolated prepared human papillomavirus  
 CC (HPV) epitope (1). (1) has antiviral activity, and can be used in  
 CC vaccine production. Peptides and corresponding nucleic acid compositions  
 CC from the present invention are useful for stimulating an immune response  
 CC to HPV by stimulating the production of CTL or HTL responses,





XX Disclosure; Fig 1b; 83bp; English.  
PS  
CC The fragment of DNA contg. the HPV-18 E6/E7 coding region was  
CC prodnd. by PCR from plasmid pBR322/HPV16 (Boshart et al., EMBO J. 3:  
CC 1151) using oligonucleotides S01 and S02. The prod. of the  
CC second reading frame is the HPV-18 E7 protein whereas the third  
CC reading frame encodes HPV-18 E6. The E6 and E7 ORFs are fused  
CC together to form a single continuous ORF via site directed mutagenesis  
CC and the immortalising potential of E7 is removed by altering two key  
CC codons of the HPV E7 sequence. The single ORF of HPV-18 E6/E7 may be  
CC inserted into vaccinia virus DNA at neutral sites (pref. by inserting  
CC two sets of the DNA in opposite orientations to overcome the problem  
CC of intertypic recombination) to make a recombinant virus vector for  
CC use immunotherapeutically to activate cells of the immune system  
CC against HPV. See also AAR27723-43.  
CC  
SQ Sequence 271 AA;

Query Match 100.0%; Score 223; DB 13; Length 271;  
Best Local Similarity 100.0%; Pred. No. 8.2e-23;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPYKLPDLCTELNLSLQDIETCYCYKTVLETFEFAFK 42  
Db 10 rrykrlpdlctelnslsqdieltcycktvleltvfeafk 51  
|||||

RESULT 10  
AAV25385  
ID AAV25385 standard; Protein; 278 AA.  
XX  
AC AAV25385;  
XX  
DT 06-SEP-1999 (first entry)  
XX  
DE HPV fusion protein D1/3-E6-His/HPV18.  
XX  
KW Fusion protein; E6 protein; E7 protein; E6/E7; Immunomodulator; tumour;  
KW Immunological fusion partner; Cpg oligonucleotide; immune response;  
KW HPV antigen; prevention; treatment.  
XX  
OS Synthetic.  
OS Human papillomavirus.  
XX  
PM WO9933868-A2.  
XX  
PD 08-JUL-1999.  
XX  
PE 18-DEC-1998; 98WO-EP08563.  
XX  
PR 24-DEC-1997; 97GB-0027262.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
PI Dalemans WLJ, Gerard CMG;  
XX  
DR WPI: 1999-405485/34.  
DR N-PSDB: AAX78800.  
XX  
PT Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to  
PT induce immune response to HPV  
XX  
PS Example XI: Page 59-60; 62pp; English.  
XX  
CC AAX78791-X78801 represent nucleic acid sequences which encode novel  
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from  
CC HPV (represented in AAY25375-Y25386). These constructs are optionally  
CC linked to an immunological fusion partner and an immunomodulatory Cpg  
CC oligonucleotide. The products of the invention can be used to induce an  
CC immune response in a patient to an HPV antigen. They can also be used  
CC for preventing or treating HPV induced tumours.

XX Sequence 278 AA;  
SQ  
Query Match 100.0%; Score 223; DB 20; Length 278;  
Best Local Similarity 100.0%; Pred. No. 8.5e-23;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPYKLPDLCTELNLSLQDIETCYCYKTVLETFEFAFK 42  
Db 120 rrykrlpdlctelnslsqdieltcycktvleltvfeafk 161  
|||||

RESULT 11  
AAV02641  
ID AAV02641 standard; Protein; 278 AA.  
XX  
AC AAV02641;  
XX  
DT 22-JUN-1999 (first entry)  
XX  
DE Prot.D1/3-E6-His/HPV18 protein.  
XX  
KW Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;  
KW tumour; lesion; benign; malignant; virus; infection.  
XX  
OS Chimeric - Human Papillomavirus.  
OS Chimeric - Haemophilus Influenzae.  
XX  
PM WO9910375-A2.  
XX  
PD 04-MAR-1999.  
XX  
PE 17-AUG-1998; 98WO-EP05285.  
XX  
PR 22-AUG-1997; 97GB-0017953.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
PI Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;  
PI Lombardo-Bencheikh A;  
XX  
DR WPI: 1999-190587/16.  
DR N-PSDB: AAX29789.  
XX  
PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for  
PT treatment or prophylaxis of HPV induced lesions  
XX  
PS Disclosure; Fig 22; 95pp; English.  
XX  
CC This sequence represents a chimeric E6 or E7 protein or E6/E7 fusion  
CC protein from Human papillomavirus (HPV) linked to an immunological fusion  
CC partner, in this case, a fragment of the Haemophilus influenzae B  
CC protein D. The sequence also contains a histidine tag at the C-terminus  
CC of the encoded protein. The protein can be used in a vaccine, for  
CC immuno-therapeutically treating HPV induced tumour lesions (benign or  
CC malignant) and preventing HPV viral infection.  
XX  
SQ Sequence 278 AA;

Query Match 100.0%; Score 223; DB 20; Length 278;  
Best Local Similarity 100.0%; Pred. No. 8.5e-23;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPYKLPDLCTELNLSLQDIETCYCYKTVLETFEFAFK 42  
Db 120 rrykrlpdlctelnslsqdieltcycktvleltvfeafk 161  
|||||

RESULT 12  
AAV25386  
ID AAV25386 standard; Protein; 383 AA.

```

XX AC AAY25386;
XX XX
XX 06-SEP-1999 (first entry)
XX DE HPV fusion protein D1/3-E6/E7-His/HPV18.
XX XX
XX fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
XX immunological fusion partner; Cpg oligonucleotide; immune response;
XX HPV antigen; prevention; treatment.
XX OS Synthetic.
XX OS Human papillomavirus.
XX PN WO933868-A2.
XX PD 08-JUL-1999.
XX PS 18-DEC-1998; 98WO-EP08563.
XX PR 24-DEC-1997; 97GB-0027262.
XX PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Dalemans WLJ, Gerard CMG;
XX DR WPI: 1999-405485/34.
XX DR N-PSDB: AAX78801.
XX PT Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
XX PT induce immune response to HPV
XX PS Example XII: Page 61-62; 62pp; English.
XX CC AAX78791-X78801 represent nucleic acid sequences which encode novel
XX CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
XX CC HPV (represented in AAY25375-Y25386). These constructs are optionally
XX CC linked to an immunological fusion partner and an immunomodulatory Cpg
XX CC oligonucleotide. The products of the invention can be used to induce an
XX CC immune response in a patient to an HPV antigen. They can also be used
XX CC for preventing or treating HPV induced tumours.
XX SQ Sequence 383 AA;

Query Match 100.0%; Score 223; DB 20; Length 383;
Best Local Similarity 100.0%; Pred. NO. 1.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPYKLPDICTELINTSLQDIEITCVYCKTVLETFEVEFAFK 42
   ||||||||||||||||||||||||||||||||||||||||
DB 120 rrpypkldictelntslqdieltcvcyckvleltevfeafk 161

RESULT 13
AAY02642
ID AAY02642 standard; Protein; 383 AA.
XX AC AAY02642;
XX XX
XX 22-JUN-1999 (first entry)
XX DE Prot.D1/3-E6-E7-His/HPV18 protein.
XX XX
XX Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
XX KM tumour; lesion; benign; malignant; virus; infection.
XX OS Chimeric - Human papillomavirus.
XX OS Chimeric - Haemophilus influenzae.
XX PN WO910375-A2.
XX PD 04-MAR-1999.

```

```

XX XX
XX 17-AUG-1998; 98WO-EP05285.
XX XX
XX 22-AUG-1997; 97GB-0017953.
XX PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
XX PI Lombardo-Bencheikh A;
XX DR WPI: 1999-190587/16.
XX DR N-PSDB: AAX29790.
XX PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
XX PT treatment or prophylaxis of HPV induced lesions
XX PS Disclosure; Fig 25; 95pp; English.
XX XX
XX This sequence represents a chimeric E6 or E7 protein or E6/E7 fusion
XX CC protein from Human papillomavirus (HPV) linked to an immunological fusion
XX CC partner. In this case, a fragment of the Haemophilus influenzae B
XX CC protein D. The sequence also contains a histidine tag at the C-terminus
XX CC of the encoded protein. The protein can be used in a vaccine, for
XX CC immuno-therapeutically treating HPV induced tumour lesions (benign or
XX CC malignant) and preventing HPV viral infection.
XX SQ Sequence 383 AA;

Query Match 100.0%; Score 223; DB 20; Length 383;
Best Local Similarity 100.0%; Pred. NO. 1.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPYKLPDICTELINTSLQDIEITCVYCKTVLETFEVEFAFK 42
   ||||||||||||||||||||||||||||||||||||||||
DB 120 rrpypkldictelntslqdieltcvcyckvleltevfeafk 161

RESULT 14
AAB98440
ID AAB98440 standard; Protein; 158 AA.
XX AC AAB98440;
XX XX
XX 22-AUG-2001 (first entry)
XX DE Human papillomavirus protein HPV45 E6.
XX XX
XX Human papillomavirus; human leukocyte antigen; HLA; immune response;
XX KM HPV; epitope; T cell; identification; vaccine; infection; genital wart;
XX KM neoplastic growth; antiviral.
XX OS Human papillomavirus.
XX PN WO200141799-A1.
XX PD 14-JUN-2001.
XX XX
XX 11-DEC-2000; 2000WO-US33549.
XX PR 10-DEC-1999; 99US-0127205.
XX PR 15-AUG-2000; 2000US-0641528.
XX PA (EPIM-) EPIMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Chesnut R, Cells E, Grey HM;
XX DR WPI: 2001-381497/40.
XX PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for
XX PT treating HPV infections -
XX PS Disclosure; Page 25; 756pp; English.

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 13, 2002, 09:59:57 ; Search time 32.4 Seconds  
(without alignments)  
29.171 Million cell updates/sec

Title: US-09-664-225-152  
Perfect score: 223  
Sequence: 1 RRRYKLPDICTELNTSLQDI.....TCVYCKTVLETFVFFFAFK 42

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 segs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCUTS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	223	100.0	158	2	US-08-247-904B-10 Sequence 10, Appl
2	223	100.0	158	3	US-08-767-942A-19 Sequence 19, Appl
3	223	100.0	271	1	US-08-117-083-14 Sequence 14, Appl
4	160	71.7	32	1	US-08-466-285-2 Sequence 2, Appl
5	130	58.3	162	1	US-08-316-239B-3 Sequence 3, Appl
6	130	58.3	162	1	US-08-316-239B-4 Sequence 4, Appl
7	130	58.3	172	3	US-08-860-165-14 Sequence 14, Appl
8	130	58.3	172	4	US-09-359-382-14 Sequence 14, Appl
9	130	58.3	182	1	US-08-117-083-10 Sequence 10, Appl
10	130	58.3	266	3	US-08-860-165-10 Sequence 10, Appl
11	130	58.3	266	4	US-09-359-382-10 Sequence 10, Appl
12	82	36.8	15	3	US-08-159-339A-1176 Sequence 1176, Ap
13	78	35.0	30	1	US-08-363-386-4 Sequence 4, Appl
14	73	32.7	15	3	US-08-159-339A-1174 Sequence 1174, Ap
15	57	25.6	376	4	US-09-461-474-6 Sequence 6, Appl
16	56	25.1	20	2	US-08-934-915-159 Sequence 159, App
17	53	23.8	10	3	US-08-159-339A-86 Sequence 86, Appl
18	50	22.4	10	3	US-08-159-339A-87 Sequence 87, Appl
19	49	22.0	102	2	US-08-808-982-8 Sequence 8, Appl
20	49	22.0	102	4	US-09-306-902A-8 Sequence 8, Appl
21	48	21.5	943	2	US-08-808-982-7 Sequence 7, Appl
22	48	21.5	943	4	US-09-306-902A-7 Sequence 7, Appl
23	47.5	21.3	695	6	5460961-5 Patent No. 5460961
24	47	21.1	9	3	US-08-159-339A-562 Sequence 562, App
25	47	21.1	84	2	US-08-605-165-22 Sequence 22, Appl
26	47	21.1	444	4	US-09-119-788-2 Sequence 2, Appl
27	47	21.1	500	1	US-07-755-573C-8 Sequence 8, Appl

28	46.5	20.9	735	4	US-09-115-704-2 Sequence 2, Appl
29	46.5	20.9	2071	3	US-09-415-522-6 Sequence 6, Appl
30	45.5	20.4	191	4	US-08-745-404-4 Sequence 4, Appl
31	45	20.2	301	2	US-08-951-871-6 Sequence 6, Appl
32	45	20.2	407	3	US-08-989-370-5 Sequence 5, Appl
33	45	20.2	1140	2	US-08-657-641-7 Sequence 7, Appl
34	45	20.2	1140	5	PCF-US94-07233-7 Sequence 7, Appl
35	44.5	20.0	334	2	US-08-449-933-6 Sequence 6, Appl
36	44.5	20.0	334	4	US-07-966-049A-6 Sequence 6, Appl
37	44.5	20.0	2818	1	US-08-510-284-1 Sequence 1, Appl
38	44.5	20.0	2818	1	US-08-411-389-2 Sequence 2, Appl
39	44.5	20.0	2818	2	US-08-449-933-2 Sequence 2, Appl
40	44.5	20.0	2818	4	US-07-966-049A-2 Sequence 2, Appl
41	44.5	20.0	2818	4	US-09-542-331-2 Sequence 2, Appl
42	44	19.7	20	2	US-08-934-915-160 Sequence 160, App
43	44	19.7	496	3	US-08-881-784-1 Sequence 1, Appl
44	44	19.7	496	4	US-09-292-768-2 Sequence 2, Appl
45	44	19.7	496	4	US-09-292-768-64 Sequence 64, Appl

## ALIGNMENTS

RESULT 1  
US-08-247-904B-10  
: Sequence 10, Application US/08247904B  
: Patent No. 5981699  
: GENERAL INFORMATION:  
: APPLICANT: Rolfe, Mark  
: APPLICANT: Eckstein, Jens W.  
: APPLICANT: Draetta, Giulio  
: TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Foley, Hoag & Elliot  
: STREET: One Post Office Square  
: CITY: Boston  
: STATE: MA  
: COUNTRY: USA  
: ZIP: 02109  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: ASCII(text)  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/247,904B  
: FILING DATE: 23-MAY-1994  
: CLASSIFICATION: 530  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Vincent, Matthew P.  
: REGISTRATION NUMBER: 36,709  
: REFERENCE/DOCKET NUMBER: MIV-029.01  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (617) 832-1000  
: TELEFAX: (617) 832-7000  
: INFORMATION FOR SEQ. ID NO. 10:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 158 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-08-247-904B-10

Query Match 100.0%; Score 223; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 7e-25;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRYKLPDICTELNTSLQDIETTCVYCKTVLETFVFFFAFK 42  
DB 9 RRRYKLPDICTELNTSLQDIETTCVYCKTVLETFVFFFAFK 50

```
RESULT 2
US-08-767-942A-19
; Sequence 19, Application US/08767942A
; Patent No. 6068982
; GENERAL INFORMATION:
; APPLICANT: Rolfe, Mark
; APPLICANT: Chiu, M. Isabel
; APPLICANT: Berlin, Vivian
; APPLICANT: Damagnez, Veronique
; APPLICANT: Draetta, Giulio
; APPLICANT: Guillaume, Cottarel
; TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/767,942A
; FILING DATE: 17-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-029,04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-767-942A-19

Query Match 100.0%; Score 223; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPYKLPDLCTELNTSLQDIEITCYCKTVLELVEFEFAFK 42
Db 9 RRPYKLPDLCTELNTSLQDIEITCYCKTVLELVEFEFAFK 50

RESULT 3
US-08-117-083-14
; Sequence 14, Application US/08117083
; Patent No. 5719054
; GENERAL INFORMATION:
; APPLICANT: Boursnell, Michael E.
; APPLICANT: Inglis, Stephen C.
; APPLICANT: Munro, Alan J.
; TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
; TITLE OF INVENTION: Papilloma Virus Proteins
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Walter H. Dreyer
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,083
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreyer, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-58783
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; TELEEX: 910 277299
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..271
; OTHER INFORMATION:
; OTHER INFORMATION: the open reading frame."
; US-08-117-083-14
```

```
Query Match 100.0%; Score 223; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.3e-24;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPYKLPDLCTELNTSLQDIEITCYCKTVLELVEFEFAFK 42
Db 10 RRPYKLPDLCTELNTSLQDIEITCYCKTVLELVEFEFAFK 51

RESULT 4
US-08-466-285-2
; Sequence 2, Application US/08466285
; Patent No. 5753233
; GENERAL INFORMATION:
; APPLICANT: Bleul, Conrad
; APPLICANT: Gissmann, Lutz
; APPLICANT: Muller, Martin
; TITLE OF INVENTION: Seroreactive Epitopes On Proteins Of
; TITLE OF INVENTION: Human Papillomavirus (HPV)18
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,285
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,768
; FILING DATE: 10-DEC-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: US 07/947,992  
FILING DATE: 21-SEP-1992  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/696,953  
FILING DATE: 08-MAY-1991  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 40 15 044.5  
FILING DATE: 10-MAY-1990  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Manspeizer, David A.  
REGISTRATION NUMBER: 37,540  
REFERENCE/DOCKET NUMBER: 05552.1075-03000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-466-285-2

Query Match 71.7%; Score 160; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.3e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPRYKLPDLCTELNTSLQDIETCYCYCTVLETFVEFPARK 29  
Db 4 RPRYKLPDLCTELNTSLQDIETCYCYCTVLETFVEFPARK 32

RESULT 5  
US-08-316-239B-3  
Sequence 3, Application US/08316239B  
Patent No. 5679509  
GENERAL INFORMATION:  
APPLICANT: Wheeler, Cosette M.  
APPLICANT: Parmenter, Cheryl A.  
TITLE OF INVENTION: Methods and a Diagnostic Aid for  
TITLE OF INVENTION: Distinguishing a Subset of HPV that is Associated with an  
TITLE OF INVENTION: Increased Risk of Developing Cervical Dysplasia and  
TITLE OF INVENTION: Cervical Cancer  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jagtiani & Associates  
STREET: 6126 Rocky Way Court  
CITY: Centreville  
STATE: VA  
COUNTRY: USA  
ZIP: 20120-3400  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,239B  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jagtiani, Ajay A.  
REGISTRATION NUMBER: 35,205  
REFERENCE/DOCKET NUMBER: UNME-0001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 817-9453  
TELEFAX: (703) 803-9387  
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 162 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-316-239B-3

Query Match 58.3%; Score 130; DB 1; Length 162;  
Best Local Similarity 58.5%; Pred. No. 1.9e-11;  
Matches 24; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 RPRYKLPDLCTELNTSLQDIETCYCYCTVLETFVEFPARK 42  
Db 15 RPRYKLPDLCTELNTSLQDIETCYCYCTVLETFVEFPARK 55

RESULT 6  
US-08-316-239B-4  
Sequence 4, Application US/08316239B  
Patent No. 5679509  
GENERAL INFORMATION:  
APPLICANT: Wheeler, Cosette M.  
APPLICANT: Parmenter, Cheryl A.  
TITLE OF INVENTION: Methods and a Diagnostic Aid for  
TITLE OF INVENTION: Distinguishing a Subset of HPV that is Associated with an  
TITLE OF INVENTION: Increased Risk of Developing Cervical Dysplasia and  
TITLE OF INVENTION: Cervical Cancer  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jagtiani & Associates  
STREET: 6126 Rocky Way Court  
CITY: Centreville  
STATE: VA  
COUNTRY: USA  
ZIP: 20120-3400  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,239B  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jagtiani, Ajay A.  
REGISTRATION NUMBER: 35,205  
REFERENCE/DOCKET NUMBER: UNME-0001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 817-9453  
TELEFAX: (703) 803-9387  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 162 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-316-239B-4

Query Match 58.3%; Score 130; DB 1; Length 162;  
Best Local Similarity 58.5%; Pred. No. 1.9e-11;  
Matches 24; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 RPRYKLPDLCTELNTSLQDIETCYCYCTVLETFVEFPARK 42  
Db 15 RPRYKLPDLCTELNTSLQDIETCYCYCTVLETFVEFPARK 55





```
? CURRENT FILING DATE: 1997-09-22
? EARLIER APPLICATION NUMBER: PCT/AU95/00868
? EARLIER FILING DATE: 1995-12-20
? EARLIER APPLICATION NUMBER: AU PNO157
? EARLIER FILING DATE: 1994-12-20
? NUMBER OF SEQ ID NOS: 15
? SOFTWARE: Patentln Ver. 2.0
? SEQ ID NO 10
? LENGTH: 266
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-10

Query Match      58.3%; Score 130; DB 3; Length 266;
Best Local Similarity 58.5%; Pred. No. 3.5e-11;
Matches 24; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY      2 RPKKLPDLCTELNTSLQDIEITCVCKTVLELVEFEFARK 42
Db      15 RPKKLPDLCTELNTSLQDIEITCVCKQQLLRREYDFAR 55

RESULT 11
US-09-359-382-10
? Sequence 10, Application US/09359382
? Patent No. 6306397
? GENERAL INFORMATION:
? APPLICANT: EDWARDS, Stirling John
? APPLICANT: COX, John Cooper
? APPLICANT: WEBB, Elizabeth Ann
? APPLICANT: FRAZER, Ian
? TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
? FILE REFERENCE: 017227/0148
? CURRENT APPLICATION NUMBER: US/09/359,382
? CURRENT FILING DATE: 1999-07-23
? EARLIER APPLICATION NUMBER: US 08/860,165
? EARLIER FILING DATE: 1997-09-22
? EARLIER APPLICATION NUMBER: PCT/AU95/00868
? EARLIER FILING DATE: 1995-12-20
? EARLIER APPLICATION NUMBER: AU PNO157/94
? EARLIER FILING DATE: 1994-12-20
? NUMBER OF SEQ ID NOS: 27
? SOFTWARE: Patentln Ver. 2.0
? SEQ ID NO 10
? LENGTH: 266
? TYPE: PRT
? ORGANISM: Human papillomavirus type 16
US-09-359-382-10

Query Match      58.3%; Score 130; DB 4; Length 266;
Best Local Similarity 58.5%; Pred. No. 3.5e-11;
Matches 24; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY      2 RPKKLPDLCTELNTSLQDIEITCVCKTVLELVEFEFARK 42
Db      15 RPKKLPDLCTELNTSLQDIEITCVCKQQLLRREYDFAR 55

RESULT 12
US-08-159-339A-1176
? Sequence 1176, Application US/08159339A
? Patent No. 6037135
? GENERAL INFORMATION:
? APPLICANT: Kubo, Ralph T.
? APPLICANT: Grey, Howard M.
? APPLICANT: Sette, Alessandro
? APPLICANT: Celis, Esteban
? TITLE OF INVENTION: HLA Binding peptides and their
? TITLE OF INVENTION: Uses
```

```
? NUMBER OF SEQUENCES: 1254
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Townsend and Townsend and Crew LLP
? STREET: Two Embarcadero Center, Eighth Floor
? CITY: San Francisco
? STATE: CA
? COUNTRY: USA
? ZIP: 94111-3834
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/159,339A
? FILING DATE: 29-NOV-1993
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/926,666
? FILING DATE: 07-AUG-1992
? APPLICATION NUMBER: US 08/027,746
? FILING DATE: 05-MAR-1993
? APPLICATION NUMBER: US 08/103,396
? FILING DATE: 06-AUG-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Weber, Ellen Lauver
? REGISTRATION NUMBER: 32,762
? REFERENCE/DOCKET NUMBER: 018623-005030US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 576-0200
? TELEFAX: (415) 576-0300
? TELEX:
? INFORMATION FOR SEQ ID NO: 1176:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 15 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
US-08-159-339A-1176

Query Match      36.8%; Score 82; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      14 NTSLDIEITCVCK 28
Db      1 NTSLDIEITCVCK 15

RESULT 13
US-08-363-586-4
? Sequence 4, Application US/08363586
? Patent No. 5629161
? GENERAL INFORMATION:
? APPLICANT: Mueller, Martin
? APPLICANT: Gissmann, Lutz
? TITLE OF INVENTION: Use of HPV-16 E6 and E7-Gene Derived
? TITLE OF INVENTION: Peptides for the Diagnostic Purpose
? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
? ADDRESSEE: Dunner
? STREET: 1300 I Street, N.W.
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005-3315
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/363,586  
FILING DATE: 23-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/909,296  
FILING DATE: 09-JUL-1992  
APPLICATION NUMBER: EP 9111720.8  
FILING DATE: 13-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Wadler, Linda A.  
REGISTRATION NUMBER: 33,218  
REFERENCE/DOCKET NUMBER: 02481-1195-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-363-586-4

Query Match 35.0%; Score 78; DB 1; Length 30;  
Best Local Similarity 60.9%; Pred. No. 8.2e-05;  
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 RPKLPDLCTELNLSLODIETTC 24  
||| ||||| : : :  
Db 8 RPKLPDLCTELNLSLODIETTC 30

RESULT 14  
US-08-159-339A-1174  
Sequence 11/4, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esteban  
TITLE OF INVENTION: HLA Binding peptides and Their  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762

REFERENCE/DOCKET NUMBER: 018623-00503005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 1174:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-159-339A-1174

Query Match 32.7%; Score 73; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00019;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 KTVLELTFEVEFPAFK 42  
||||| ||||| |||||  
Db 1 KTVLELTFEVEFPAFK 15

RESULT 15  
US-09-461-474-6  
Sequence 6, Application US/09461474  
Patent No. 6278042  
GENERAL INFORMATION:  
APPLICANT: Allen, Steve  
APPLICANT: Rafalski, Antoni  
APPLICANT: Sakai, Hajime  
TITLE OF INVENTION: Plant Metal Transporters  
FILE REFERENCE: B81303 US NA  
CURRENT APPLICATION NUMBER: US/09/461,474  
CURRENT FILING DATE: 1999-12-14  
EARLIER APPLICATION NUMBER: 60/112,562  
EARLIER FILING DATE: 1998-12-16  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 6  
LENGTH: 376  
TYPE: PRT  
ORGANISM: Trifolium aestivum  
US-09-461-474-6

Query Match 25.6%; Score 57; DB 4; Length 376;  
Best Local Similarity 37.0%; Pred. No. 1.9; Mismatches 10; Conservative 5; Indels 0; Gaps 0;  
Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

OY 12 ELNLSLODIETTCVYCKTVLELTFEVE 38  
||| ||| : : :  
Db 260 EVNTOFQDAELTFVCVCAELSLYE 286

Search completed: February 13, 2002, 09:59:58  
Job time: 347 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 10:00:43 ; Search time 39.34 Seconds

(without alignments)  
81.325 Million cell updates/sec

Title: US-09-664-225-152

Perfect score: 223  
Sequence: 1 RRPYKLPDLCTELNTSLQDI.....TCVYCKTVLELTVFFFAFK 42

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	223	100.0	158	1 M6WL18	E6 protein - human
2	191	85.7	57	2 S36705	E6 protein - human
3	187	83.9	158	2 S36561	E6 protein - human
4	161	72.2	158	1 M6WL39	E6 protein - human
5	148	66.4	158	1 M6WLPR	E6 protein - human
6	130	58.3	158	1 M6WLHS	E6 protein - human
7	127	57.0	148	2 S36515	E6 protein - human
8	120	53.8	149	1 M6WL35	E6 protein - human
9	113	50.7	150	2 S36544	E6 protein - human
10	106	47.5	149	1 M6WL58	E6 protein - human
11	105	47.1	149	1 M6WL33	E6 protein - human
12	100	44.8	154	2 S36555	E6 protein - human
13	97	43.5	151	1 M6WL51	E6 protein - human
14	92	41.3	153	2 S36503	E6 protein - human
15	91	40.8	149	1 M6WL31	E6 protein - human
16	91	40.8	154	2 S36527	E6 protein - human
17	90	40.4	150	1 M6WL44	E6 protein - human
18	90	40.4	154	2 S36584	E6 protein - human
19	89	39.9	155	1 M6WL43	E6 protein - human
20	88.5	39.7	150	1 M6WL13	E6 protein - human
21	88	39.5	148	2 A61237	E6 protein - human
22	88	39.5	148	2 S36573	E6 protein - human
23	85	38.1	150	1 M6WL6	E6 protein - human
24	83	37.2	150	1 M6WL11	E6 protein - human
25	83	36.8	150	1 S15621	E6 protein - pygmy
26	82	36.3	153	1 S15621	E6 protein - human
27	79	35.4	150	1 M6WL42	E6 protein - human
28	79	35.4	155	1 M6WL56	E6 protein - human
29	78	35.0	142	2 S36509	E6 protein - human

30	78	35.0	191	1 M6WL1
31	77	34.5	159	1 S15614
32	77	34.5	159	2 S36497
33	76	34.1	155	2 A44890
34	72	32.3	32	2 S19906
35	72	32.3	35	2 S19909
36	72	32.3	47	2 S23825
37	67	30.0	138	2 S36567
38	66.5	29.8	676	2 T48104
39	63	28.3	155	1 M6WL8
40	63	28.3	156	1 M6WL41
41	62.5	28.0	154	2 S28509
42	62	27.8	148	2 S36532
43	59	26.5	152	2 S36550
44	58	26.0	660	2 T22794
45	58	26.0	2358	2 T39569

#### ALIGNMENTS

```

RESULT 1
M6WL18
E6 protein - human papillomavirus type 18
C:Species: human papillomavirus type 18
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: A26165; G26251
R:Seedorf, K.; Oltersdorf, T.; Kraemer, G.; Roewekamp, W.
EMBO J. 6, 139-144, 1987
A:Title: Identification of early proteins of the human papilloma viruses type 16 (HPV
A:Reference number: A91068; M01D:87218459
A:Accession: A26165
A:Molecule type: DNA
A:Residues: 1-158 <SEP>
A:Cross-References: GB:X04773; NID:q60876; PIDN:CAA28466.1; PID:q60877
R:Cole, S.T.; Danos, O.
J. Mol. Biol. 193, 599-608, 1987
A:Title: Nucleotide sequence and comparative analysis of the human papillomavirus typ
A:Reference number: A92937; M01D:87283882
A:Accession: G26251
A:Molecule type: DNA
A:Residues: 1-158 <COL>
A:Cross-References: GB:X05015; NID:q60975; PIDN:CAA28664.1; PID:q60976
R:Mathiasewski, G.; Banks, L.; Wu-Liao, J.; Spence, P.; Pim, D.; Crawford, L.
J. Gen. Virol. 67, 1909-1916, 1986
A:Title: The expression of human papillomavirus type 18 E6 protein in bacteria and th
A:Reference number: A92791; M01D:86306665
A:Contents: annotation; Identification of the protein
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; transforming protein; zinc finger
F:32-68/Region: zinc finger CCCC motif
F:105-141/Region: zinc finger CCCC motif

Query Match 100.0%; Score 223; DB 1; Length 158;
Best local Similarity 100.0%; Pred. No. 6.6e-21;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPYKLPDLCTELNTSLQDIETCVYCKTVLELTVFFFAFK 42
DB 9 RRPYKLPDLCTELNTSLQDIETCVYCKTVLELTVFFFAFK 50

RESULT 2
I56705
E6 protein - human papillomavirus type 18
C:Species: human papillomavirus type 18
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 26-Aug-1999
C:Accession: I56705
R:Inagaki, Y.; Tanokawa, Y.; Takebe, N.; Nawa, H.; Nakanishi, S.; Terada, M.; Sugimu
J. Virol. 62, 1640-1646, 1988
A:Title: Nucleotide sequences of cDNAs for human papillomavirus type 18 transcripts 1
A:Reference number: I56705; M01D:88188247

```

A:Accession: I56705  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-57 <RES>  
A:Cross-references: GB:M0324; NID:q183933; PIDN:AAA9512.1; PID:q306834  
C:Superfamily: papillomavirus E6 protein

Query Match 85.7%; Score 191; DB 2; Length 57;  
Best Local Similarity 100.0%; Pred. No. 2.7e-17;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPYKLPDCLTELNTSLQDIETTCYCKTVLETFEFAF 36  
|||||  
DB 9 RRPYKLPDCLTELNTSLQDIETTCYCKTVLETFEFAF 44

## RESULT 3

S36561  
E6 protein - human papillomavirus type 45  
C:Species: human papillomavirus type 45  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S36561  
R:Delius, H.; Hofmann, B.  
submitted to the EMBL Data Library, August 1993  
A:Description: primer-directed sequencing of human papillomavirus types.  
A:Reference number: S36469  
A:Accession: S36561  
A:Molecule type: DNA  
A:Residues: 1-158 <DEL>  
A:Cross-references: EMBL:X74479; NID:q397022; PIDN:CAA52573.1; PID:q397023  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 83.9%; Score 187; DB 2; Length 158;  
Best Local Similarity 78.6%; Pred. No. 2.2e-16;  
Matches 33; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRPYKLPDCLTELNTSLQDIETTCYCKTVLETFEFAF 42  
|||||  
DB 9 RRPYKLPDCLTELNTSLQDVSIACVCKATLETFEYGFAR 50

## RESULT 4

M6WL39  
E6 protein - human papillomavirus type 39  
C:Species: human papillomavirus type 39  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999  
C:Accession: A38502  
R:Volpers, C.; Strebeck, R.E.  
Virology 181, 419-423, 1991  
A:Title: Genome organization and nucleotide sequence of human papillomavirus type 39.  
A:Reference number: A38502; MUID:91135017  
A:Accession: A38502  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-158 <VOL>  
A:Cross-references: GB:M62849; EMBL:M38185; NID:q333245; PIDN:AAA47050.1; PID:q463186  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; transforming protein; zinc finger  
F:32-68/Region: zinc finger CCCC motif  
F:105-141/Region: zinc finger CCCC motif

Query Match 72.2%; Score 161; DB 1; Length 158;  
Best Local Similarity 72.5%; Pred. No. 4.1e-13;  
Matches 29; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 RPYKLPDCLTELNTSLQDIETTCYCKTVLETFEFAF 41  
|||||  
DB -10 RPYKLPDCLTLDITLQDIETTCYCKTVLETFEFAF 49

RESULT 5  
M6MLPR  
E6 protein - human papillomavirus type ME180 (provirus)

C:Species: human papillomavirus type ME180  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Feb-1997  
C:Accession: C40509  
R:Reuter, S.; Delius, H.; Kahn, T.; Hofmann, B.; zur Hausen, H.; Schwarz, E.  
J. Virol. 65, 5564-5568, 1991  
A:Title: Characterization of a novel human papillomavirus DNA in the cervical carcinoma cell line C40509.  
A:Reference number: A40509; MUID:91374616  
A:Accession: C40509  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-158 <REU>

A:Cross-references: GB:M73258  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; zinc finger  
F:32-68/Region: zinc finger CCCC motif  
F:105-141/Region: zinc finger CCCC motif

Query Match 66.4%; Score 148; DB 1; Length 158;  
Best Local Similarity 65.0%; Pred. No. 1.8e-11;  
Matches 26; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 2 RPYKLPDCLTELNTSLQDIETTCYCKTVLETFEFAF 41  
|||||  
DB 10 RPYKLPDCLTLDITLQDIETTCYCKTVLETFEFAF 49

## RESULT 6

M6WL39  
E6 protein - human papillomavirus type 16  
C:Species: human papillomavirus type 16  
C:Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 20-Aug-1999  
C:Accession: A03682; T10427  
R:Seedorf, K.; Krammer, G.; Durst, M.; Suhai, S.; Roweckamp, W.G.  
Virology 145, 181-185, 1985  
A:Title: Human papillomavirus type 16 DNA sequence.  
A:Reference number: A22355; MUID:85246220  
A:Accession: A03682  
A:Molecule type: DNA  
A:Residues: 1-158 <SEB>  
A:Cross-references: GB:K02718; NID:q333031; PIDN:AAA46939.1; PID:q333032  
J. Kennedy, I.M.; Haddow, J.K.; Clements, J.B.  
J. Virol. 65, 2093-2097, 1991  
A:Title: A negative element in the human poapillomavirus type 16 genome acts at the 1  
A:Reference number: Z17014; MUID:91162763  
A:Accession: T10427  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-158 <KEN>  
A:Cross-references: EMBL:K02718; NID:q333031; PIDN:AAA46939.1; PID:q333032  
C:Genetics:  
A:Gene: E6  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; zinc finger  
F:37-73/Region: zinc finger CCCC motif  
F:110-146/Region: zinc finger CCCC motif

Query Match 58.3%; Score 130; DB 1; Length 158;  
Best Local Similarity 58.5%; Pred. No. 3.3e-09;  
Matches 24; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 RPYKLPDCLTELNTSLQDIETTCYCKTVLETFEFAF 42  
|||||  
DB 15 RPYKLPDCLTLDITLQDIETTCYCKTVLETFEFAF 55

```
RESULT 7
S36515
E6 protein - human papillomavirus type 34
C:Species: human papillomavirus type 34
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36515
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36515
A:Molecule type: DNA
A:Residues: 1-148 <DEL>
A:Cross-references: EMBL:X74476; NID:g396989; PIDN:CAA52555.1; PID:g396990
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 57.0%; Score 127; DB 2; Length 148;
Best Local Similarity 53.7%; Pred. No. 7.4e-09;
Matches 22; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

OY 2 RPYKLPDICTELNTSLQDIEITGVCKTVLETFEVEFAK 42
Db 9 RPYKLPALCEVNIHIEIETDCYCEQLYRCEVYDFIFR 49

RESULT 8
W6WL35
E6 protein - human papillomavirus type 35
C:Species: human papillomavirus type 35
A:Note: host Homo sapiens (man)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Mar-2001
C:Accession: E40824; S36521
R:Marich, J.E.; Pontsler, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.
Virology 186, 770-776, 1992
A>Title: The phylogenetic relationship and complete nucleotide sequence of human papillomavirus type 35
A:Reference number: A40824; MUID:92124753
A:Accession: E40824
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-149 <MAR>
A:Cross-references: GB:M74117; NID:g333050; PIDN:AAA46966.1; PID:g333051
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <DEL>
A:Cross-references: EMBL:X74477; NID:g396997; PIDN:CAA52561.1; PID:g396998
A:Experimental source: strain 35H
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; zinc finger
F:30-66/Region: zinc finger CCCC motif
F:103-139/Region: zinc finger CCCC motif

Query Match 53.8%; Score 120; DB 1; Length 149;
Best Local Similarity 53.8%; Pred. No. 5.6e-08;
Matches 21; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

OY 2 RPYKLPDICTELNTSLQDIEITGVCKTVLETFEVEFA 40
Db 8 RPYKLPALCEVNIHIEIETDCYCEQLYRCEVYDFIFR 46

RESULT 9
S36544
E6 protein - human papillomavirus type 26
C:Species: human papillomavirus type 26
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
```

```
C:Accession: S36544
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36544
A:Molecule type: DNA
A:Residues: 1-150 <DEL>
A:Cross-references: EMBL:X74472; NID:g396956; PIDN:CAA52530.1; PID:g396957
C:Superfamily: papillomavirus E6 protein
C:Keywords: early protein; zinc finger

Query Match 50.7%; Score 113; DB 2; Length 150;
Best Local Similarity 48.7%; Pred. No. 4.3e-07;
Matches 19; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

OY 2 RPYKLPDICTELNTSLQDIEITGVCKTVLETFEVEFA 40
Db 8 RPYKLPALCEVNIHIEIETDCYCEQLYRCEVYDFIFR 46

RESULT 10
W6WL58
E6 protein - human papillomavirus type 58
C:Species: human papillomavirus type 58
A:Note: host Homo sapiens (man)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
C:Accession: E36779
R:Kiril, Y.; Iwamoto, S.; Matukura, T.
Virology 185, 424-427, 1991
A>Title: Human papillomavirus type 58 DNA sequence.
A:Reference number: A36779; MUID:92024102
A:Accession: E36779
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-149 <KIR>
A:Cross-references: GB:D90400; NID:g722386; PIDN:BAA31845.1; PID:g3337098
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; zinc finger
F:30-66/Region: zinc finger CCCC motif
F:103-139/Region: zinc finger CCCC motif

Query Match 47.5%; Score 106; DB 1; Length 149;
Best Local Similarity 47.5%; Pred. No. 3.2e-06;
Matches 19; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

OY 2 RPYKLPDICTELNTSLQDIEITGVCKTVLETFEVEFA 41
Db 8 RPYKLPALCEVNIHIEIETDCYCEQLYRCEVYDFIFR 47

RESULT 11
W6WL33
E6 protein - human papillomavirus type 33
C:Species: human papillomavirus type 33
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03683
R:Cole, S.T.; Strecek, R.E.
J. Virol. 58, 991-995, 1986
A>Title: Genome organization and nucleotide sequence of human papillomavirus type 33,
A:Reference number: A93020; MUID:86200464
A:Accession: A03683
A:Molecule type: DNA
A:Residues: 1-149 <COL>
A:Cross-references: GB:M12732; NID:g333049; PIDN:AAA46958.1; PID:g463177
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; zinc finger
F:30-66/Region: zinc finger CCCC motif
F:103-139/Region: zinc finger CCCC motif
```

Query Match 47.1%; Score 105; DB 1; Length 149;  
Best Local Similarity 47.5%; Pred. No. 4.3e-06;  
Matches 19; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

OY 2 RPRKLPDCLTELNTSLQDIETVCYCKTVLELTFEVEFAF 41  
DB 8 RPRTLHDLCOALETTIHNIIELOCVCCKRPLORSEVYDPAF 47

## RESULT 12

S36555

E6 protein - human papillomavirus type 40

C:Species: human papillomavirus type 40

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36555

R:Deilus, H.; Hofmann, B.

Submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36555

A:Molecule type: DNA

A:Residues: 1-154 &lt;DEL&gt;

A:Cross-references: EMBL:X74478; NID:q397014; PIDN:CAA52567.1; PID:q671878

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 44.8%; Score 100; DB 2; Length 154;

Best Local Similarity 48.6%; Pred. No. 1.9e-05;

Matches 18; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

OY 6 LPDCLTELNTSLQDIETVCYCKTVLELTFEVEFAF 42  
DB 12 LVELCDCCNITLPTLOIDCVCKTVLKTAEVLAFAFR 48

## RESULT 13

W6WLS1

E6 protein - human papillomavirus type 51

C:Species: human papillomavirus type 51

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997

C:Accession: E40415

R:Lungu, O.; Crum, C.P.; Silverstein, S.J.

A:Title: Biologic properties and nucleotide sequence analysis of human papillomavirus ty

A:Reference number: A40415; MUID:91303675

A:Accession: E40415

A&gt;Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-151 &lt;LUN&gt;

A:Cross-references: GB:M62877

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; zinc finger

F:30-66/Region: zinc finger CCCC motif

F:103-139/Region: zinc finger CCCC motif

Query Match 43.5%; Score 97; DB 1; Length 151;  
Best Local Similarity 45.0%; Pred. No. 4.5e-05;  
Matches 18; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

OY 2 RPRKLPDCLTELNTSLQDIETVCYCKTVLELTFEVEFAF 41  
DB 8 RPRTLHELCEALNYSMHNIQVCYCKKELCRADYVNAF 47

## RESULT 14

S36503

E6 protein - human papillomavirus type 30

C:Species: human papillomavirus type 30

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36503

R:Deilus, H.; Hofmann, B.  
submitted to the EMBL Data Library, August 1993  
A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36503

A:Molecule type: DNA

A:Residues: 1-153 &lt;DEL&gt;

A:Cross-references: EMBL:X74474; NID:q396973; PIDN:CAA52543.1; PID:q396974

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 41.3%; Score 92; DB 2; Length 153;  
Best Local Similarity 43.9%; Pred. No. 0.00019;  
Matches 18; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

OY 2 RPRKLPDCLTELNTSLQDIETVCYCKTVLELTFEVEFAF 42  
DB 11 RPRVHDLCEVQETSLTELQCVCKKELSSSEVYNFAF 51

## RESULT 15

W6WLS1

E6 protein - human papillomavirus type 31

C:Species: human papillomavirus type 31

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 16-Jul-1999

C:Accession: A32444

R:Goldsbrough, M.D.; Disilvestre, D.; Temple, G.F.; Lorincz, A.T.

Virology 171, 306-311, 1989

A:Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-as

A:Reference number: A94398; MUID:89299478

A:Accession: A32444

A&gt;Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-149 &lt;GOL&gt;

A:Cross-references: GB:J04353; NID:q333048; PIDN:AAA46950.1; PID:q459916

C:Comment: This protein may be involved in the oncogenic potential of this virus.

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; zinc finger

F:30-66/Region: zinc finger CCCC motif

F:103-139/Region: zinc finger CCCC motif

Query Match 40.8%; Score 91; DB 1; Length 149;  
Best Local Similarity 45.0%; Pred. No. 0.00025;  
Matches 18; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

OY 2 RPRKLPDCLTELNTSLQDIETVCYCKTVLELTFEVEFAF 41  
DB 8 RPRKLHLSALEIPYDELRLNCVCCKGOLTFETEVIDPAF 47

Search completed: February 13, 2002, 10:00:44  
Job time: 378 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 10:02:20 ; Search time 21.6 seconds

(without alignments)  
71.293 Million cell updates/sec

Title: US-09-664-225-152

Perfect score: 223  
Sequence: 1 RRPYKLPDLCTELNTSLQDI.....TCVYCKIVLELVEFEFAFK 42

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	223	100.0	158	1	VE6_HPVI8
2	187	83.9	158	1	VE6_HPVI8
3	164	73.5	158	1	VE6_HPVI8
4	161	72.2	158	1	VE6_HPVI8
5	148	66.4	158	1	VE6_HPVI8
6	148	66.4	158	1	VE6_HPVI8
7	130	58.3	158	1	VE6_HPVI8
8	127	57.0	148	1	VE6_HPVI8
9	120	53.8	148	1	VE6_HPVI8
10	113	50.7	150	1	VE6_HPVI8
11	106	47.5	150	1	VE6_HPVI8
12	105	47.1	149	1	VE6_HPVI8
13	100	44.8	154	1	VE6_HPVI8
14	97	43.5	151	1	VE6_HPVI8
15	92	41.3	153	1	VE6_HPVI8
16	91	40.8	149	1	VE6_HPVI8
17	91	40.8	154	1	VE6_HPVI8
18	90	40.4	150	1	VE6_HPVI8
19	90	40.4	154	1	VE6_HPVI8
20	89	39.9	155	1	VE6_HPVI8
21	88.5	39.7	150	1	VE6_HPVI8
22	88	39.5	148	1	VE6_HPVI8
23	88	39.5	150	1	VE6_HPVI8
24	88	39.5	150	1	VE6_HPVI8
25	85	38.1	150	1	VE6_HPVI8
26	83	37.2	150	1	VE6_HPVI8
27	82	36.8	153	1	VE6_HPVI8
28	81	36.3	150	1	VE6_HPVI8
29	80	35.9	140	1	VE6_HPVI8
30	79	35.4	146	1	VE6_HPVI8
31	79	35.4	155	1	VE6_HPVI8
32	78	35.0	140	1	VE6_HPVI8
33	78	35.0	142	1	VE6_HPVI8

# ALIGNMENTS

RESULT	1	STANDARD:	PRT:	158 AA.
VE6_HPVI8	1			
ID	VE6_HPVI8			
AC	P06463			
DT	01-JAN-1988 (Rel. 06, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	E6 PROTEIN.			
GN	E6.			
OS	Human papillomavirus type 18.			
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			
OC	Papillomavirus.			
OX	NCHI_Taxid=10582;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=67283882; PubMed=3039146;			
RA	Cole S.T., Danos O.;			
RT	"Nucleotide sequence and comparative analysis of the human			
RT	papillomavirus type 18 genome. Phylogeny of papillomaviruses and			
RT	repeated structure of the E6 and E7 gene products.";			
RL	J. Mol. Biol. 193:599-608(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=6630665; PubMed=3018129;			
RA	Matlashewski G., Banks L., Wu-Liao J., Spence P., Plm D., Crawford L.;			
RT	"The expression of human papillomavirus type 18 E6 protein in			
RT	bacteria and the production of anti-E6 antibodies.";			
RT	J. Gen. Virol. 67:1909-1916(1986).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88188247; PubMed=2833614;			
RA	Inagaki Y., Tsunokawa Y., Takebe N., Nawa H., Nakanishi S.;			
RT	"Nucleotide sequences of cDNAs for human papillomavirus type 18			
RT	transcripts in HeLa cells.";			
RL	J. Virol. 62:1640-1646(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87053870; PubMed=3023067;			
RA	Schneider-Gaedtke A., Schwarz E.;			
RT	"Different human cervical carcinoma cell lines show similar			
RT	transcription patterns of human papillomavirus type 18 early genes.";			
RL	EMBO J. 5:2285-2292(1986).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87218459; PubMed=3034571;			
RA	Seedorf K., Oltersdorf T., Kraemer G., Roewekamp W.;			
RT	"Identification of early proteins of the human papilloma viruses type			
RT	16 (HPV 16) and type 18 (HPV 18) in cervical carcinoma cells.";			
RL	EMBO J. 6:139-144(1987).			
RN	[6]			
RP	ZINC-BINDING.			
RX	MEDLINE=89385606; PubMed=2550872;			
RA	Grossman S.R., Laimins L.A.;			
RT	"E6 protein of human papillomavirus type 18 binds zinc.";			
RT	Oncogene 4:1089-1093(1989).			

34	78	35.0	144	1	VE6_HPVI8	O81018 human papil
35	78	35.0	191	1	VE6_HPVI8	P22159 rhesus papil
36	77	34.5	153	1	VE6_HPVI8	P25484 human papil
37	77	34.5	159	1	VE6_HPVI8	P36808 human papil
38	76	34.1	155	1	VE6_HPVI8	O80955 human papil
39	72	32.3	142	1	VE6_HPVI8	O80941 human papil
40	71	31.8	146	1	VE6_HPVI8	P50802 human papil
41	67	30.0	138	1	VE6_HPVI8	P36813 human papil
42	66	29.6	140	1	VE6_HPVI8	P50777 human papil
43	65	29.1	144	1	VE6_HPVI8	O89808 human papil
44	63	28.3	155	1	VE6_HPVI8	P06428 human papil
45	63	28.3	156	1	VE6_HPVI8	P27555 human papil

```

CC -1- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -----
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X04354; CAA27879.1; -
CC DR EMBL; X05015; CAA28864.1; -
CC DR EMBL; M20325; AAA9514.1; -
CC DR EMBL; M26798; AAA46946.1; -
CC DR EMBL; X04773; CAA28466.1; -
CC DR EMBL; A06324; CAA00539.1; -
CC DR EMBL; A06328; CAA00542.1; -
CC DR PIR; G26251; W6W1D8.
CC DR InterPro: IPR001334; E6.
CC DR Pfam; PF00518; E6; 1.
CC KW Early protein: DNA-binding; Nuclear protein; Zinc-finger;
CC Transforming protein.
CC KW ZN_FING 32 68 POTENTIAL.
CC FT ZN_FING 105 141 POTENTIAL.
CC FT CONFLICT 22 22 N -> S (IN REF. 4).
CC FT SEQUENCE 158 AA; 18871 MW; 5BCF13CEFA3D157FA CRC64;
CC -----
OY * 1 RRPYKLPDLCETELNLSODIEITYCYCKTVLE50EVFEPAFK 42
Db 9 RRPYKLPDLCETELNLSODIEITYCYCKTVLETFEVEFAFK 50
-----
RESULT 2
VE6_HPVA5 STANDARD; PRT: 158 AA.
AC P21735;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE E6 PROTEIN.
GN E6..
OS Human papillomavirus type 45.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10593;
NN [1]
NN SEQUENCE FROM N.A.
RX MEDLINE-9426501; PubMed-8205838;
RA Delius R., Hofmann B.;
RT *Primer-directed sequencing of human papillomavirus types.*;
RN Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Kaplan J.B., Burk R.D.;
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -----
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74479; CAA52573.1; -.
DR EMBL; M38198; AAA46973.1; -.
DR PIR; S36561; S36561.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
DR Early protein: DNA-binding; Nuclear protein: Zinc-finger;
KW Transforming protein.
FT ZN_FING 32 68
FT ZN_FING 105 141 POTENTIAL.
FT CONFLICT 10 10 R -> P (IN REF. 2).
FT CONFLICT 30 30 I -> N (IN REF. 2).
FT CONFLICT 118 118 R -> A (IN REF. 2).
SO SEQUENCE 158 AA; 18897 MW; F1CF10DD33AA4C3E CRC64;

Query Match 83.9%; Score 187; DR 1; Length 158;
Best Local Similarity 78.6%; Pred. No. 2,3e-16;
Matches 33; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 RRPYKLPDLCTELMTSLQDIEETCYCKVLETFEVEFAFK 42
Db 9 GRPYKLPDLCTELMTSLQDVSIACYCATLERFVEYFAFK 50

RESULT 3
VE6_HPV70
ID VE6_HPV70 STANDARD; PRT; 156 AA.
AC P50804;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 70.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=39457;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96249586; PubMed=8015087;
RA Forslund O., Hansson B.G.;
RT "Human papillomavirus type 70 genome cloned from overlapping PCR
RT products: complete nucleotide sequence and genomic organization.";
RL J. Clin. Microbiol. 34:802-809(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97060129; PubMed=8904450;
RA Longuet M., Beaudenon S., Orth G.;
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,
RT related to the potentially oncogenic HPV39.";
RL J. Clin. Microbiol. 34:738-744(1996).
CC -I- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -I- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U21941; AAC54850.1; -.
DR EMBL; U22461; AAC54880.1; -.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
DR Early protein: DNA-binding; Nuclear protein: Zinc-finger.
FT ZN_FING 32 68 POTENTIAL.
FT ZN_FING 105 141 POTENTIAL.

```



FT CONFLICT 100 100 N -> D (IN REF. 2).  
SQ SEQUENCE 158 AA; 18617 MW; 6B610800D923D6DE CRC64;

Query Match 73.5%; Score 164; DB 1; Length 158;  
Best Local Similarity 75.0%; Pred. No. 3e-15;  
Matches 30; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 2 RPKLPDLCTELNTSLQDIEITCVYCKTVLELVEFEFAR 41  
DB 10 RPKLPDLCTALDTLDTLIDVTIDCVYCKTQLOQTEVEFEFAR 49

RESULT 4  
VE6\_HPV39 STANDARD; PRT; 158 AA.

AC P24635;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE E6 PROTEIN.  
GN E6.  
OS Human papillomavirus type 39.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10586;

RP SEQUENCE FROM N.A.  
RA MEDLINE-91135017; PubMed-1847266;  
RA Volpers C., Streck R.E.;  
RT Genome organization and nucleotide sequence of human papillomavirus  
RT type 39.";  
RL Virology 181:419-423(1991).

CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE ONCOGENIC POTENTIAL  
OF THIS VIRUS (CERVICAL NEOPLASIA-ASSOCIATED VIRUS).

CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
STRANDED DNA (IN VITRO).

CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.

CC "This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
DR EMBL: M62849; AAA47050.1; -.  
DR PIR: A38502; W6WL39.  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1  
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;  
KM Transferring protein.  
FT ZN\_FING 32 68 POTENTIAL.  
FT ZN\_FING 105 141 POTENTIAL.  
SQ SEQUENCE 158 AA; 18726 MW; 1B5E9D5BC1B662E CRC64;

Query Match 72.2%; Score 161; DB 1; Length 158;  
Best Local Similarity 72.5%; Pred. No. 7.5e-15;  
Matches 29; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 2 RPKLPDLCTELNTSLQDIEITCVYCKTVLELVEFEFAR 41  
DB 10 RPKLPDLCTALDTLDTLIDVTIDCVYCKRPLQOTEVEFEFAR 49

RESULT 5

VE6\_HPV68 STANDARD; PRT; 158 AA.

AC P54667;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE E6 PROTEIN.  
GN E6.

OS Human papillomavirus type 68.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=45240;

RP SEQUENCE FROM N.A.  
RA MEDLINE-97060129; PubMed-8904450;  
RA Longuet M., Beaudenon S., Orth G.;  
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,  
RT related to the potentially oncogenic HPV39.";  
RL J. Clin. Microbiol. 34:738-744(1996).  
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
STRANDED DNA (IN VITRO).

CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: X67160; CAA47632.1; -.  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1.  
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
FT ZN\_FING 32 68 POTENTIAL.  
FT ZN\_FING 105 141 POTENTIAL.  
SQ SEQUENCE 158 AA; 18796 MW; 46B37939CBA6596 CRC64;

Query Match 66.4%; Score 148; DB 1; Length 158;  
Best Local Similarity 65.0%; Pred. No. 4.3e-13;  
Matches 26; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 2 RPKLPDLCTELNTSLQDIEITCVYCKTVLELVEFEFAR 41  
DB 10 RPKLPDLCTALDTLDTLIDVTIDCVYCKRQLORETEVEFEFAR 49

RESULT 6

VE6\_HPV6E STANDARD; PRT; 158 AA.

AC P27962;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE E6 PROTEIN.

GN Human papillomavirus type ME180.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10602;

RP SEQUENCE FROM N.A.  
RA MEDLINE-91374616; PubMed-1716694;  
RA Reuter S., Delius H., Kahn T., Hofmann B., Zur Hausen H., Schwarz E.;  
RT "Characterization of a novel human papillomavirus DNA in the cervical  
RT carcinoma cell line ME180.";  
RL J. Virol. 65:5564-5568(1991).

CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
STRANDED DNA (IN VITRO).

CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: M73258; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: C40509; W6WLP; .  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
 FT ZN-FING 32 68 POTENTIAL.  
 FT ZN-FING 106 141 POTENTIAL.  
 SO SEQUENCE 158 AA; 18738 MW; 2B1F40B563F05FC CRC64;

Query Match Best Local Similarity 66.4%; Score 148; DB 1; Length 158;  
 Matches 26; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 2 RPKLPDLCETLNTSLQDIETCYCKVLELTFEYEFAR 41  
 Db 10 RPKLPDLCRTLDPTLHDVTIDCYCRQLQRTVEYEFAR 49

## RESULT 7

VE6\_HPV16 STANDARD; PRT: 158 AA.  
 AC P03126;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE E6 PROTEIN.  
 GN E6.  
 OS Human Papillomavirus type 16.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10581;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=85246220; PubMed=2990099;  
 RT Seedorf K., Kramer G., Durst M., Suhai S., Roweckamp W.G.;  
 RT "Human papillomavirus type 16 DNA sequence.";  
 RT Virology 145:181-185(1985).  
 RN 121

RP SEQUENCE OF 31-50 FROM N.A.  
 RX MEDLINE=90218027; PubMed=2157796;  
 RA Schneider-Maunoury S., Pehau-Arnaudet G., Breitburd F., Orth G.;  
 RT "Expression of the human papillomavirus type 16 genome in SK-V cells,  
 RT a line derived from a vulvar intraepithelial neoplasia.";  
 RT J. Gen. Virol. 71:809-817(1990)  
 CC -1- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.  
 CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
 CC -1- STRANDED DNA (IN VITRO).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.  
 CC -1- MISCELLANEOUS: HPV16, IN COMPARISON TO HPV TYPES 6 AND 11, IS MORE  
 CC OFTEN ASSOCIATED WITH MALIGNANT GENITAL CANCERS IN HUMANS.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: K02718; AAA46939.1; -;  
 DR EMBL: D00735; BAA00632.1; -;

DR PIR: A03682; W6WLP; .  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.

KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;  
 KM Transforming protein.

FT ZN-FING 37 73 POTENTIAL.  
 FT ZN-FING 110 146 POTENTIAL.

SO SEQUENCE 158 AA; 19187 MW; 01FEF5A0CFDB37EB CRC64;

Query Match Best Local Similarity 58.3%; Score 130; DB 1; Length 158;  
 Matches 24; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 2 RPKLPDLCETLNTSLQDIETCYCKVLELTFEYEFAR 42  
 Db 15 RPKLPDLCETLQTTIHDLIECYCKQQLRREYDEAFR 55

## RESULT 8

VE6\_HPV34 STANDARD; PRT: 148 AA.  
 AC P36811;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE E6 PROTEIN.  
 GN E6.  
 OS Human Papillomavirus type 34.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10613;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94265501; PubMed=8205838;  
 RA Delius H., Holmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types.";  
 RT Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
 CC -1- STRANDED DNA (IN VITRO).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X74476; CA52555.1; -;  
 DR PIR: S36515; S36515.  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
 FT ZN-FING 31 67 POTENTIAL.  
 FT ZN-FING 104 140 POTENTIAL.  
 SO SEQUENCE 148 AA; 17735 MW; E2FC6E62E4AF0DA CRC64;

Query Match Best Local Similarity 57.0%; Score 127; DB 1; Length 148;  
 Matches 22; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

OY 2 RPKLPDLCETLNTSLQDIETCYCKVLELTFEYEFAR 42  
 Db 9 RPKLPALCEVNSIHIEILDYCCERQLRCEYDIFR 49

## RESULT 9

VE6\_HPV35 STANDARD; PRT: 149 AA.  
 AC P27228;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE E6 PROTEIN.  
 GN E6.  
 OS Human Papillomavirus type 35.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.

```

OX NCB1_Taxid=10587;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-94265501; PubMed-8205838;
RA Dellus H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE-92124753; PubMed-1310198;
RA Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of
human papillomavirus type 35.";
RL Virology 186:770-776(1992).
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE ONCOGENIC POTENTIAL
CC OF THIS VIRUS (ASSOCIATED WITH CANCER OF THE UTERINE CERVIX).
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X74477; CAAG52561.1; -
CC DR EMBL: M74117; AAA6966.1; -
CC DR PIR: E40824; W6WL55.
CC DR PIR: S36521; S36521.
CC DR InterPro: IPR001334; E6.
CC DR Pfam: PF00518; E6; 1.
CC KM Early protein: DNA-binding; Nuclear protein; Zinc-finger;
CC KM Transforming protein.
CC FT ZN_FING 103 139 POTENTIAL.
CC FT ZN_FING 103 139 POTENTIAL.
CC SO SEQUENCE 149 AA; 18045 MW; C605D19AF3935021 CRC64;

Query Match 53.8%; Score 120; DB 1; Length 149;
Best Local Similarity 53.8%; Pred. No. 2.5e-09;
Matches 21; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 2 RPKKLPDLCTELNTSLQDIEITGVCKTVLETFEVEFA 40
DB 8 RPKKHLDCLENTSVEISHEICLNCVCKEQLQRESEVDF 46

RESULT 10
VE6_HPV26 STANDARD; PRT; 150 AA.
ID VE6_HPV26 STANDARD; PRT; 150 AA.
AC P36807;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 26.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_Taxid=31549;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-94265501; PubMed-8205838;
RA Dellus H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X74472; CAAG52530.1; -
CC DR PIR: S36544; S36344.
CC DR InterPro: IPR001334; E6.
CC DR Pfam: PF00518; E6; 1.
CC SO SEQUENCE 150 AA; 17923 MW; BAA6C9459750AB15 CRC64;

Query Match 50.7%; Score 113; DB 1; Length 150;
Best Local Similarity 48.7%; Pred. No. 2.2e-08;
Matches 19; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 2 RPKKLPDLCTELNTSLQDIEITGVCKTVLETFEVEFA 40
DB 8 RPKKHLDCLENTSLTQNLQVCKEQLQRESEVDF 46

RESULT 11
VE6_HPV58 STANDARD; PRT; 149 AA.
ID VE6_HPV58 STANDARD; PRT; 149 AA.
AC P26355;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 58.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCB1_Taxid=10598;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-92024102; PubMed-1656594;
RA Kiril Y., Iwamoto S., Matsukura T.;
RT "Human papillomavirus type 58 DNA sequence.";
RL Virology 185:424-427(1991).
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D90400; BAA31845.1; -
CC DR PIR: E36779; W6WL58.
CC DR InterPro: IPR001334; E6.
CC DR Pfam: PF00518; E6; 1.
CC KM Early protein: DNA-binding; Nuclear protein; Zinc-finger.
CC FT ZN_FING 30 66 POTENTIAL.
CC FT ZN_FING 103 139 POTENTIAL.
CC SO SEQUENCE 149 AA; 17794 MW; 79B3DC95831B158 CRC64;

Query Match 47.5%; Score 106; DB 1; Length 149;
Best Local Similarity 47.5%; Pred. No. 2e-07;
Matches 19; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 2 RPKKLPDLCTELNTSLQDIEITGVCKTVLETFEVEFA 41
DB 8 RPKKHLDCLENTSVEISHEICLNCVCKEQLQRESEVDF 47

```

```

RESULT 12
VE6_HP33 STANDARD: PRT: 149 AA.
ID VE6_HP33
AC P06427:
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 33.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10586;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86200464; PubMed=3009902;
RA Cole S.T., Streeck R.E.;
RT "Genome organization and nucleotide sequence of human papillomavirus
RT type 33, which is associated with cervical cancer.";
RL J. Virol. 58:991-995(1986).
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE ONCOGENIC POTENTIAL
CC OF THIS VIRUS (CERVICAL NEOPLASIA-ASSOCIATED VIRUS).
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed.. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M12732; AAA46958.1; -.
DR PIR: A03683; W6M133.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;
KW Transforming protein.
FT ZN_FING 30 66 POTENTIAL.
FT ZN_FING 103 139 POTENTIAL.
SQ SEQUENCE 149 AA; 17652 MW; EDD87444F9C8B1AF CRC64;

Query Match 47.1%; Score 105; DB 1; Length 149;
Best Local Similarity 47.5%; Pred. No. 2.7e-07;
Matches 19; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

OY 2 RPYKLPDLCTELNLSLQDIETCYVCKTVLETFEVEFAF 41
DB 8 KPRTHDLCQALLETTHINLELCVCKRPLRSEVYDAF 47

RESULT 13
VE6_HP40 STANDARD: PRT: 154 AA.
ID VE6_HP40
AC P36812:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 40.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";

```

```

RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed.. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74478; CAAS2567.1; -.
DR PIR: S36555; S36555.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT ZN_FING 30 66 POTENTIAL.
FT ZN_FING 103 139 POTENTIAL.
SQ SEQUENCE 154 AA; 17894 MW; 54C90FE52002F5CD CRC64;

Query Match 44.8%; Score 100; DB 1; Length 154;
Best Local Similarity 48.6%; Pred. No. 1.3e-06;
Matches 18; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

OY 6 LPDLCTELNLSLQDIETCYVCKTVLETFEVEFAF 42
DB 12 LXEIDQCNTLPTLQIDCVCKTVLKTAEVLAFAFR 48

RESULT 14
VE6_HP51 STANDARD: PRT: 151 AA.
ID VE6_HP51
AC P26554:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 51.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10595;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91303675; PubMed=1649326;
RA Lungu O., Crum C.P., Silverstein S.J.;
RT "Biologic properties and nucleotide sequence analysis of human
RT papillomavirus type 51."
RL J. Virol. 65:4216-4225(1991).
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed.. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M62877; -. NOT_ANNOTATED_CDS.
DR PIR: E40415; W6M151.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT ZN_FING 30 66 POTENTIAL.
FT ZN_FING 103 139 POTENTIAL.
SQ SEQUENCE 151 AA; 18134 MW; 577E663BA1376478 CRC64;

```

Query Match 43.5%; Score 97; DB 1; Length 151;  
 Best Local Similarity 45.0%; Pred. No. 3.3e-06;  
 Matches 18; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

OY 2 RPYKLPDLCTELNTSLQDIEITFCVCKVLELTFEVEFAF 41  
 || | : || || | : || : || || | : || : ||  
 DB 8 RPYKLPDLCTELNTSLQDIEITFCVCKVLELTFEVEFAF 47

## RESULT 15

VE6\_HPV30 STANDARD; PRT; 153 AA.  
 AC P36809;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE E6 PROTEIN.  
 GN E6.  
 OS Human papillomavirus type 30.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus  
 OX NCBI\_TaxId=10611;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94265501; PubMed-8205838;  
 RA Delius H., Hofmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types."  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
 CC STRANDED DNA (IN VITRO).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.  
 CC  
 CC \*This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC  
 CC  
 DR EMBL: X74474; CAAS2543.1; -.  
 DR PIR: S36503; S36503.  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
 FT ZN\_FING 33 69 POTENTIAL.  
 FT ZN\_FING 106 142 POTENTIAL.  
 SQ SEQUENCE 153 AA; 18266 MW; F7295E15261FC433 CRC64;

Query Match 41.3%; Score 92; DB 1; Length 153;  
 Best Local Similarity 43.9%; Pred. No. 1.6e-05;  
 Matches 18; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

OY 2 RPYKLPDLCTELNTSLQDIEITFCVCKVLELTFEVEFAF 42  
 || | : || || | : || : || || | : || : ||  
 DB 11 RPYKLPDLCTELNTSLQDIEITFCVCKVLELTFEVEFAF 51

Search completed: February 13, 2002, 10:02:20  
 Job time: 249 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 13, 2002, 10:01:53 ; Search time 62.52 Seconds  
(without alignments)  
98.264 Million cell updates/sec

Title: US-09-664-225-152  
Perfect score: 223  
Sequence: 1 RRPYKLPDICTELNTSLQDI.....TCVYCKVLELFEVFEFAFK 42

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPTRMBL\_17:\*  
2: sp.archaea:\*  
3: sp.bacteria:\*  
4: sp.fungi:\*  
5: sp.human:\*  
6: sp.invertebrate:\*  
7: sp.mammal:\*  
8: sp.mhc:\*  
9: sp.organelle:\*  
10: sp.phage:\*  
11: sp.plant:\*  
12: sp.rodent:\*  
13: sp.virus:\*  
14: sp.vertibrate:\*  
15: sp.unclassified:\*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	223	100.0	158	12	Q9QNP8 human papill
2	191	85.7	157	12	Q9Q133 human papill
3	187	83.9	158	4	Q9Y4Y4 homo sapien
4	187	83.9	158	12	Q10608 human papill
5	177	79.4	157	12	Q9WHG0 human papill
6	148	66.4	162	4	Q9Q129 Q9uiz2 homo sapien
7	139	62.3	151	12	Q81964 human papill
8	137	61.4	151	12	Q89708 human papill
9	136	61.0	151	12	Q12335 human papill
10	136	61.0	151	12	Q12336 human papill
11	136	61.0	151	12	Q80966 human papill
12	132	59.2	151	12	Q9WMP3 human papill
13	132	59.2	151	12	Q9WMP2 human papill
14	130	58.3	151	12	Q80963 human papill
15	130	58.3	151	12	Q9WMP5 human papill
16	130	58.3	151	12	Q9WMP4 human papill
17	130	58.3	151	12	Q9WBC3 human papill
18	130	58.3	151	12	Q89755 human papill
19	130	58.3	151	12	Q89852 human papill

20	130	58.3	158	12	Q9WH13 human papill
21	129	57.8	151	12	Q89887 human papill
22	129	57.8	151	12	Q89648 human papill
23	129	57.8	151	12	Q9W931 human papill
24	129	57.8	158	12	Q9QDH7 human papill
25	129	57.8	158	12	Q9QDH5 human papill
26	126	56.5	151	12	Q89640 human papill
27	126	56.5	158	12	Q9QDH9 human papill
28	126	56.5	158	12	Q9QDH3 human papill
29	120	53.8	149	12	Q84298 human papill
30	118	52.9	148	12	Q82005 human papill
31	110	49.3	151	12	Q9JH51 human papill
32	106	47.5	149	12	Q9JH51 human papill
33	106	47.5	149	12	Q9JH51 human papill
34	96	43.0	151	12	Q9JH51 human papill
35	95	42.6	151	12	Q9JH51 human papill
36	94	42.2	149	12	Q9JH51 human papill
37	90	40.4	150	12	Q9JH51 human papill
38	87	39.0	150	12	Q80934 human papill
39	84	37.7	148	12	Q81997 human papill
40	82	36.8	147	12	Q9WMP5 human papill
41	73.5	33.0	152	12	Q9JH51 human papill
42	72	32.3	35	12	Q81885 human papill
43	72	32.3	35	12	Q81887 human papill
44	72	32.3	47	12	Q81883 human papill
45	72	32.3	47	12	Q81884 human papill

## ALIGNMENTS

RESULT 1	Q9QNP8	PRELIMINARY;	PRT;	158 AA.
AC	Q9QNP8;			
DT	01-MAY-2000 (Tremblrel, 13, Created)			
DT	01-MAY-2000 (Tremblrel, 13, Last sequence update)			
DT	01-JUN-2001 (Tremblrel, 17, Last annotation update)			
DE	E6 PROTEIN.			
GN	E6.			
OS	Human papillomavirus type 18.			
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			
OC	Papillomavirus.			
OX	NCBI_TaxID=10562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Laasari M., Gul'ko L., Vinokurova S., Kissel'ova N., Velko V.,			
RA	Kissel'ev F.;			
RT	*Cloning of E6 and E7 Genes of Human Papilloma Virus Type 18 and			
RT	Transformation Potential of E7 Gene and its Mutants.;			
RL	Virus Genes 182:139-149(1999).			
DR	EMBL; Y18491; CAB53096.1; -			
DR	InterPro; IPR001334; E6.			
DR	Pfam; PF00518; E6; 1.			
SQ	SEQUENCE 158 AA; 18886 MW; 5BCF13CF43D407AF CRC64;			
Query Match	100.0%;	Score 223;	DB 12;	Length 158;
Best local Similarity	100.0%;	Pred. No. 1.1e-22;		
Matches 42;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
QY	1 RRPYKLPDICTELNTSLQDIETCVYCKVLELFEVFEFAFK 42			
DB	9 RRPYKLPDICTELNTSLQDIETCVYCKVLELFEVFEFAFK 50			
RESULT 2	Q9Q133	PRELIMINARY;	PRT;	57 AA.
AC	Q9Q133;			
DT	01-NOV-1996 (Tremblrel, 01, Created)			
DT	01-NOV-1996 (Tremblrel, 01, Last sequence update)			
DT	01-NOV-1998 (Tremblrel, 08, Last annotation update)			

DE PROTEIN E6.  
 OS Human Papillomavirus type 18.  
 OC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10582;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=88188247; PubMed=2833614;  
 RA Inagaki Y., Tsunokawa Y., Takebe N., Nawa H., Nakinishi S., Terada M.,  
 RA Sugimura T.;  
 RT "Nucleotide sequences of cDNAs for human papillomavirus type 18  
 RL J. Virol. 62:1640-1646(1988).  
 DR EMBL; M20324; AAA9512.1; -;  
 SQ SEQUENCE 57 AA; 6531 MW; 31CF1A65B3740D2C CRC64;

Query Match 85.7%; Score 191; DB 12; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-19;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPYKLPDCLTELNTSLQDIETCYCKTVLETFEYFARK 36  
 ID 9 RRPYKLPDCLTELNTSLQDIETCYCKTVLETFEY 44

RESULT 3  
 OY4Y4 PRELIMINARY; PRT; 158 AA.  
 ID O9Y4Y4  
 AC O9Y4Y4;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE E6 PROTEIN.  
 GN HPV45 E6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Sastre-Garau X., Favre M., Couturier J., Orth G.;  
 RT "Distinct patterns of alterations of myc genes associated with  
 RT integration of HPV16 or HPV45 dna in two genital tumors."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ242956; CAB44706.1; -;  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 DR PIR: P00518; E6; 1.  
 SQ SEQUENCE 158 AA; 18914 MW; A61AEF98390AEBB3 CRC64;

Query Match 83.9%; Score 187; DB 4; Length 158;  
 Best Local Similarity 78.6%; Pred. No. 8.3e-18;  
 Matches 33; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 RRPYKLPDCLTELNTSLQDIETCYCKTVLETFEYFARK 42  
 ID 9 RRPYKLPDCLTELNTSLQDIETCYCKTVLETFEYFARK 50

RESULT 4  
 OI0608 PRELIMINARY; PRT; 158 AA.  
 ID OI0608  
 AC OI0608;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE ONCOPROTEIN E6.  
 GN E6.  
 OS Human Papillomavirus type 45.  
 CC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10593;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-IC4 HPV45 VARIANT;  
 RA Sastre-Garau X., Favre M., Couturier J., Orth G.;  
 RL J. Gen. Virol. 0:0-0(0);  
 DR EMBL; Y13218; CAA73660.1; -;  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 SQ SEQUENCE 158 AA; 18914 MW; A61AEF98390AEBB3 CRC64;

Query Match 83.9%; Score 187; DB 12; Length 158;  
 Best Local Similarity 78.6%; Pred. No. 8.3e-18;  
 Matches 33; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 RRPYKLPDCLTELNTSLQDIETCYCKTVLETFEYFARK 42  
 ID 9 RRPYKLPDCLTELNTSLQDIETCYCKTVLETFEYFARK 50

RESULT 5  
 O9WHG0 PRELIMINARY; PRT; 157 AA.  
 ID O9WHG0  
 AC O9WHG0;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE PUTATIVE TRANSFORMING PROTEIN E6.  
 GN E6.  
 OS Human papillomavirus.  
 CC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10566;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC SPRAIN-HL77474-S;  
 RA Chow Y.T.K., Leong W.F.;  
 RT "Complete nucleotide sequence of a novel genital human papillomavirus  
 RT H177474-S from Singapore."  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF131950; AAD24181.1; -;  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 DR PIR: P00518; E6; 1.  
 SQ SEQUENCE 157 AA; 18450 MW; 1B154BA2C49EDED7 CRC64;

Query Match 79.4%; Score 177; DB 12; Length 157;  
 Best Local Similarity 80.0%; Pred. No. 1.9e-16;  
 Matches 32; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 RPYKLPDCLTELNTSLQDIETCYCKTVLETFEYFARK 41  
 ID 10 RPYKLPDCLTELNTSLQDIETCYCKTVLETFEYFARK 49

RESULT 6  
 Q9UI29 PRELIMINARY; PRT; 162 AA.  
 ID Q9UI29  
 AC Q9UI29;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE HUMAN CELLULAR DNA/HUMAN PAPILLOMAVIRUS PROVIRAL DNA.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE-CERVIX;  
 RC MEDLINE=91374616; PubMed=1716694;  
 RA Reuter S., Delius H., Kahn T., Hofmann B., Zur Hausen H., Schwarz E.;  
 RT "Characterization of a novel human papillomavirus DNA in the cervical





```
RESULT 10
012336 ID 012336 PRELIMINARY; PRT; 151 AA.
AC 012336:
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE E6 PROTEIN.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
NCBI_TaxID=10581;
RN NCBI_TaxID=10581;
RP SEQUENCE FROM N.A.
RC STRAIN=PC-7;
RA MEDLINE=97437474; PubMed=9292007;
RA Tornesello M.L., Buonaguro F.M., Meglio A., Buonaguro L.,
RA Beth-Giraldo E., Giraldo G.,
RT "Sequence variations and viral genomic state of human papillomavirus
RT type 16 in penile carcinomas from Ugandan patients.";
RL J. Gen. Virol. 78:0-0(0).
DR EMBL: AF003016; AAB70733.1; -
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18206 MW; 51C12A8B149C6D9A CRC64;

Query Match 61.0%; Score 136; DB 12; Length 151;
Best Local Similarity 61.0%; Pred. No. 6.8e-11;
Matches 25: Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 2 RPKLPDCTELTSLQDIETCYVCKTVLETFEVEFAFK 42
DB 8 RPKLPDCTELQTTIHDIIECVYCKQQLRREYDFAFR 48

RESULT 11
080966 ID 080966 PRELIMINARY; PRT; 151 AA.
AC 080966: 012925; 012650; 012926; 012651; 012927; 012652; 080962;
AC 080964; 080965;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE EARLY TRANSFORMING PROTEIN E6.
OS Human papillomavirus, and
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
NCBI_TaxID=10566, 10581;
RN NCBI_TaxID=10566, 10581;
RP SEQUENCE FROM N.A.
RC SPECIES=Human papillomavirus; STRAIN=VARIOUS STRAINS;
RA MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jensen S.A.;
RT "Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments.";
RL J. Virol. 69:7743-7753(1995).
RN 12;
RP SEQUENCE FROM N.A.
RC SPECIES=Human papillomavirus; STRAIN=VARIOUS STRAINS;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN 13;
RP SEQUENCE FROM N.A.
RC SPECIES=Human papillomavirus type 16;
RA Tornesello M.L., Buonaguro F.M., Meglio A., Buonaguro L.,
RA Beth-Giraldo E., Giraldo G.,
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
```

```

[4]
RN SEQUENCE FROM N.A.
RC SPECIES=Human papillomavirus type 16; STRAIN=AFG7611T;
RA van duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;
RT "Analysis of Human papillomavirus type 16 E6 variants in relation to
RT p53 codon 72 polymorphism genotype in cervical carcinogenesis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003014; AAB70731.1; -
DR EMBL: AF003017; AAB70734.1; -
DR EMBL: U34131; AAA91678.1; -
DR EMBL: AF003013; AAB70730.1; -
DR EMBL: U34114; AAA91661.1; -
DR EMBL: U34125; AAA91672.1; -
DR EMBL: U34130; AAA91677.1; -
DR EMBL: AJ388068; CAB45128.1; -
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18292 MW; 35012A9E01993C35 CRC64;

Query Match 61.0%; Score 136; DB 12; Length 151;
Best Local Similarity 61.0%; Pred. No. 6.8e-11;
Matches 25: Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 2 RPKLPDCTELTSLQDIETCYVCKTVLETFEVEFAFK 42
DB 8 RPKLPDCTELQTTIHDIIECVYCKQQLRREYDFAFR 48

RESULT 12
Q9WMP3 ID Q9WMP3 PRELIMINARY; PRT; 151 AA.
AC Q9WMP3:
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE E6 PROTEIN.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
NCBI_TaxID=10581;
RN NCBI_TaxID=10581;
RP SEQUENCE FROM N.A.
RC STRAIN=AA/AF;
RA van duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;
RT "Analysis of Human papillomavirus type 16 E6 variants in relation to
RT p53 codon 72 polymorphism genotype in cervical carcinogenesis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ388065; CAB45122.1; -
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18412 MW; E794A494F0DE8209 CRC64;

Query Match 59.2%; Score 132; DB 12; Length 151;
Best Local Similarity 58.5%; Pred. No. 2.4e-10;
Matches 24: Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 RPKLPDCTELTSLQDIETCYVCKTVLETFEVEFAFK 42
DB 8 RPKLPDCTELQTTIHDIIECVYCKQQLRREYDFAFR 48

RESULT 13
Q9WMP2 ID Q9WMP2 PRELIMINARY; PRT; 151 AA.
AC Q9WMP2:
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
```

```

DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF077997;
RA van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Waldoemers J.M.;
RT *Analysis of Human papillomavirus type 16 E6 variants in relation to
RT p53 codon 72 polymorphism genotype in cervical carcinogenesis.";
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ386067; CAB45126.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18306 MW; 5E71BAEFA493C34 CRC64;

```

```

Query Match 59.2%; Score 132; DB 12; Length 151;
Best Local Similarity 58.5%; Pred. No. 2.4e-10;
Matches 24; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

```

```

OY 2 RPKLPDLCTELNTSLQDIEITCVCKTVLELTFEVEFAFK 42
DB 8 RPKLPDLCTELNTSLQDIEITCVCKTVLELTFEVEFAFK 48

```

```

RESULT 14
O080963 PRELIMINARY: PRT; 151 AA.
AC O080963:
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE EARLY TRANSFORMING PROTEIN E6.
GN E6.
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 4997;
RX MEDLINE=96079021; PubMed=7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jenison S.A.;
RT *Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments.";
RL J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OR 4997;
RA Farmer A.D.;
RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL: U34122; AAA91669.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18291 MW; 97C7028D5169382D CRC64;

```

```

Query Match 58.3%; Score 130; DB 12; Length 151;
Best Local Similarity 58.5%; Pred. No. 4.4e-10;
Matches 24; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

```

```

OY 2 RPKLPDLCTELNTSLQDIEITCVCKTVLELTFEVEFAFK 42
DB 8 RPKLPDLCTELNTSLQDIEITCVCKTVLELTFEVEFAFK 48

```

RESULT 15

```

O9WMP5 PRELIMINARY: PRT; 151 AA.
AC O9WMP5:
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ET182G;
RA van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Waldoemers J.M.;
RT *Analysis of Human papillomavirus type 16 E6 variants in relation to
RT p53 codon 72 polymorphism genotype in cervical carcinogenesis.";
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ242681; CAB45381.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18320 MW; 617D2D5FD93F8917 CRC64;

```

```

Query Match 58.3%; Score 130; DB 12; Length 151;
Best Local Similarity 58.5%; Pred. No. 4.4e-10;
Matches 24; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

```

```

OY 2 RPKLPDLCTELNTSLQDIEITCVCKTVLELTFEVEFAFK 42
DB 8 RPKLPDLCTELNTSLQDIEITCVCKTVLELTFEVEFAFK 48

```

Search completed: February 13, 2002, 10:01.54  
Job time: 253 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 09:59:18 : Search time 66.5 Seconds  
(without alignments)  
18.936 Million cell updates/sec

Title: US-09-664-225-69

Perfect score: 92

Sequence: 1 LLMGTLLGIVCPICISQKP 17

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.Geneseq-1101:\*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT:\*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	17	22	AA895962
2	92	100.0	21	12	AA815572
3	92	100.0	28	20	AAW93291
4	92	100.0	30	12	AA815576
5	92	100.0	98	13	AA822767
6	92	100.0	98	14	AA82361
7	92	100.0	98	19	AAW46886
8	92	100.0	98	20	AAW08020
9	92	100.0	98	21	AAV58474
10	92	100.0	98	21	AAV57721
11	92	100.0	98	21	AAV50703

12	92	100.0	98	22	AA866332
13	92	100.0	98	22	AA898421
14	92	100.0	98	22	AAU01718
15	92	100.0	98	22	AA867546
16	92	100.0	98	22	AA831607
17	92	100.0	98	22	AA849453
18	92	100.0	117	22	AA896050
19	92	100.0	121	22	AA831608
20	92	100.0	150	21	AAV50704
21	92	100.0	150	21	AAV50705
22	92	100.0	172	17	AA897563
23	92	100.0	185	20	AAW99370
24	92	100.0	198	22	AA831616
25	92	100.0	220	20	AAV25375
26	92	100.0	220	20	AAV25378
27	92	100.0	220	20	AAV02834
28	92	100.0	220	20	AAV02631
29	92	100.0	236	22	AA896052
30	92	100.0	239	20	AAV25380
31	92	100.0	239	20	AAV02636
32	92	100.0	253	20	AAV41132
33	92	100.0	253	20	AAV43480
34	92	100.0	253	20	AAV01502
35	92	100.0	253	20	AAW97612
36	92	100.0	253	20	AAW87562
37	92	100.0	253	20	AAW81586
38	92	100.0	262	13	AA827724
39	92	100.0	266	17	AA897561
40	92	100.0	288	22	AAU02129
41	92	100.0	295	22	AA831615
42	92	100.0	324	22	AA831613
43	92	100.0	371	20	AAV25377
44	92	100.0	371	20	AAV02633
45	92	100.0	390	20	AAV25381

#### ALIGNMENTS

RESULT 1	
AA895962	standard; Peptide; 17 AA.
XX	XX
AC	AA895962;
XX	XX
DT	25-JUN-2001 (first entry)
XX	XX
DE	HPV 16 E7 protein fragment SEQ ID 69.
KW	Epitope; tumour antigen; antiviral; immunostimulatory; cervical cancer;
KW	Human papillomavirus-associated disease; condyloma; cervical dysplasia;
KW	cervical dysplasia; major histocompatibility complex; MHC I.
OS	
OS	Human papillomavirus.
PN	WO200119408-A1.
XX	XX
PD	22-MAR-2001.
XX	XX
PF	18-SEP-2000; 2000WO-US25559.
XX	XX
PR	16-SEP-1999; 99US-0154665.
PR	16-SEP-1999; 99US-0398534.
PR	09-DEC-1999; 99US-0169846.
PR	09-DEC-1999; 99US-0458173.
XX	XX
PA	(ZYCO-) ZYCO5 INC.
PI	Hedley ML, Urban RC, Chicz RM.
XX	XX
DR	WPI; 2001-265996/27.
XX	XX
PT	Novel nucleic acids encoding polypeptide polypeptides containing

PT multiple epitopes from one or more proteins, useful for treating tumors  
PT and as vaccines against pathogenic agents -  
PS Claim 42; Page 52; 64pp; English.  
XX  
CC This invention relates to polynucleotides encoding a hybrid polypeptide  
CC comprising a signal sequence and three segments that are either  
CC contiguous or separated by a spacer amino acid or spacer peptide. The  
CC invention specifically details polynucleotides encoding a polypeptide  
CC peptide where the peptide segments are tumour antigens or a naturally  
CC occurring protein of a pathogenic agent. The polypeptide peptides exhibit  
CC antiviral and immunostimulatory activity. The polypeptide and  
CC polypeptide peptides are useful for eliciting an immune response in a  
CC mammal. The polynucleotide and protein are useful as vaccines for  
CC treating tumours and pathogenic infections. The polynucleotide is also  
CC useful for preventing or treating human papillomavirus (HPV)-associated  
CC diseases, particularly exophytic condyloma, flat condyloma, cervical  
CC cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV  
CC infection, cervical dysplasia, high grade squamous intraepithelial  
CC lesions, and anal HPV infection. The polynucleotide and polypeptide are  
CC useful for generating or enhancing prophylactic or therapeutic immune  
CC response against pathogens, tumours or autoimmune diseases in a  
CC population of individuals having diverse MHC allotypes, as positive  
CC controls in T cell stimulation assays in vitro, and as tools to  
CC understand processing of epitopes within cells. Peptides  
CC AAB93894 - AAB96037 and AAB96044 - AAB96048 represent major  
CC histocompatibility complex I (MHC I) associated tumour and pathogen  
CC antigens. The peptides can be used as part of the polypeptide proteins of  
CC the invention. Also included are examples of the polypeptide proteins  
CC represented by AAB96050 - AAB96052, and localisation signal peptides  
CC AAB96038 - AAB96043 and AAB96049 which can be used in the construction of  
CC the polypeptide peptides.  
XX  
SQ Sequence 17 AA:  
  
Query Match 100.0%; Score 92; DB 22; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.9e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLMGTLGIVCPICSGKP 17  
Db 1 LLMGTLGIVCPICSGKP 17  
  
RESULT 2  
AAR15572  
ID AAR15572 standard; Protein: 21 AA.  
XX  
AC AAR15572;  
XX  
DT 02-MAR-1992 (first entry)  
XX  
DE Immunopeptide #5 derived from HPV16 E7 peptide.  
XX  
KW cervical cancer: cervical intraepithelial neoplasia; CIN; wart;  
XX squamous cell carcinoma; ELISA; HPV 16.  
XX  
OS Synthetic.  
XX  
PN WO9118294-A.  
XX  
PD 28-NOV-1991.  
XX  
PF 13-MAY-1991; 91WO-SF00335.  
XX  
PR 11-MAY-1990; 90SE-0001705.  
XX  
PA (MEDS-) MEDSCAND AB.  
XX  
PI Dillner J, Dillner L, Cheng HM;  
XX  
DR WPI: 1991-369390/50.

XX  
PT Diagnosis of human papilloma virus infection and PV-carrying  
PT tumours - using synthetic peptide(s) to detect virus specific  
PT antigen-antibody complexes by immunoassay  
XX  
PS Disclosure; Page 38; 72pp; English.  
XX  
CC This is one of two peptides which have been synthesised on the  
CC basis of the amino acid sequence for the E7 protein of HPV 16. The  
CC selection of peptide sequences was based on the assumption that an  
CC immunoreactive region might be situated in the same relative region  
CC of a protein from different HPV types. The peptides were used in  
CC diagnostic immunoassays to detect HPV-infection.  
CC See AAR15523-R15601.  
XX  
SQ Sequence 21 AA:  
  
Query Match 100.0%; Score 92; DB 12; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.6e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLMGTLGIVCPICSGKP 17  
Db 5 LLMGTLGIVCPICSGKP 21  
  
RESULT 3  
AAW93291  
ID AAW93291 standard; peptide: 28 AA.  
XX  
AC AAW93291;  
XX  
DT 27-MAY-1999 (first entry)  
XX  
DE Human papillomavirus peptide fragment #5.  
XX  
KW Diagnosis: skin; immune reaction; onco-protein; E6; E7.  
XX  
OS Human papillomavirus.  
XX  
FN DE19737409-A1.  
XX  
PD 04-MAR-1999.  
XX  
PF 27-AUG-1997; 97DE-1037409.  
XX  
PR 27-AUG-1997; 97DE-1037409.  
XX  
PA (MEDI-) MEDIGENE AG.  
XX  
PI Hoepfl R;  
XX  
DR WPI: 1999-168276/15.  
XX  
DE  
XX  
PT Diagnosis kit for testing skin for immune reactions against  
PT onco-protein E6 and E7 - comprises onco-protein E6 and E7 and/or  
PT immunologically active parts of E6 and E7 derived from human  
PT papilloma virus  
XX  
PS Disclosure; Column 3; 4pp; German.  
XX  
CC This invention describes peptides used in a diagnosis kit for testing  
CC skin for immune reactions against onco-protein E6 and E7. The method  
CC of the invention comprises onco-protein E6 and E7 and/or immunologically  
CC active parts of E6 and E7 derived from human papilloma virus.  
XX  
SQ Sequence 28 AA:  
  
Query Match 100.0%; Score 92; DB 20; Length 28;  
Best Local Similarity 100.0%; Pred. No. 4.8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMGTLGIVCPICSOXP 17  
 ||||||||||||||||  
 DB 12 LMGTLGIVCPICSOXP 28

## RESULT 4

AAR15576  
 ID AAR15576 standard; Protein; 30 AA.

AC AAR15576;

DT 02-MAR-1992 (first entry)

DE Immunopeptide #9 derived from HPV16 E7 peptide.

KW cervical cancer; cervical intraepithelial neoplasia; CIN; wart;  
 squamous cell carcinoma; ELISA; HPV 16.

OS Synthetic.

PN WO9118294-A.

PD 28-NOV-1991.

PE 13-MAY-1991; 91WO-SE00335.

PR 11-MAY-1990; 90SE-0001705.

PA (MEDS-) MEDSCAND AB.

PI Dillner J, Dillner L, Cheng HM;

DR WPI: 1991-369390/50.

PT Diagnosis of human papilloma virus infection and PV-carrying  
 tumours - using synthetic peptide(s) to detect virus specific  
 antigen-antibody complexes by immunoassay

PS Disclosure: Page 38; 72pp; English.

CC This is one of two peptides which have been synthesised on the  
 basis of the amino acid sequence for the E7 protein of HPV 16. The  
 CC selection of peptide sequences was based on the assumption that an  
 CC immunoreactive region might be situated in the same relative region  
 CC of a protein from different HPV types. The peptides were used in  
 CC diagnostic immunoassays to detect HPV-infection.  
 CC See AAR15523-R15601.

SQ Sequence 30 AA;

Query Match 100.0%; Score 92; DB 12; Length 30;

Best Local Similarity 100.0%; Pred. No. 5.2e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMGTLGIVCPICSOXP 17  
 ||||||||||||||||  
 DB 14 LMGTLGIVCPICSOXP 30

## RESULT 5

AAR22767  
 ID AAR22767 standard; peptide; 98 AA.

AC AAR22767;

DT 21-SEP-1992 (first entry)

DE HPV E7 peptide.

KW Human; papillomavirus; immunogenic; cervical; warts; carcinoma;  
 cancer.

XX Synthetic.  
 OS Homo sapiens.

PN WO9205248-A.

PD 02-APR-1992.

PE 26-SEP-1991; 91WO-US07081.

PR 26-SEP-1990; 90US-0588384.

PA (BRIM ) BRISTOL-MYERS SQUIB.

PI Blake J, Chen L, Hellstrom I, Hellstrom K, Hu S L;

DR WPI: 1992-132119/16.

PT Immunogenic peptide(s) derived from E6 or E7 region of HPV16 -  
 and recombinant cells encoding them, useful in treatment and  
 PT prophylaxis of cervical warts or cancer resulting from HPV  
 PT infection

PS Disclosure: Fig 7; 81pp; English.

CC The peptide is the sequence of the human papillomavirus HPV 16 E7  
 CC nucleoprotein. Peptides corresponding to regions (prel. epitopic  
 CC regions) of HPV 16 E7 were synthesised by standard Merrifield  
 CC synthesis. Examples of such peptides are E7 1-10, 29-50 or 70-81.  
 CC Compositions contg. these peptides, antibodies against the peptides,  
 CC or recombinant cells contg. the gene encoding the immunogenic  
 CC peptides may be utilised in methods for inhibiting and treating HPV  
 CC infection and tumour initiation and progression e.g. in the  
 CC prevention or retardation of cervical warts and cervical carcinoma  
 CC resulting from HPV infection.  
 CC See also AAR22766.

SQ Sequence 98 AA;

Query Match 100.0%; Score 92; DB 13; Length 98;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMGTLGIVCPICSOXP 17  
 ||||||||||||||||  
 DB 82 LMGTLGIVCPICSOXP 98

## RESULT 6

AAR42361  
 ID AAR42361 standard; protein; 98 AA.

AC AAR42361;

DT 21-MAY-1994 (first entry)

DE Human papillomavirus 16 E7 protein and fragments.

DE Tumours; cows; horses; donkeys; regression; udder warts; HPV16.

OS Synthetic.

PN WO9320844-A.

PD 28-OCT-1993.

PE 01-APR-1993; 93WO-GB00679.

PR 08-APR-1992; 92GB-0007701.

PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

```

XX Campo MS;
PI
XX WPI; 1993-351368/44.
DR
XX Use of papilloma-virus E7 protein or fragments for the therapy of
PT papilloma-virus disease - for the regression of tumours e.g.
PT removal of warts from udders or mouth of milking cows or for
PT treatment of horses or donkeys
XX
PS Disclosure: Fig 2; 31pp; English.
XX
CC The sequence is that of the human papillomavirus type 16 E7 protein.
CC The protein sequence was aligned with that of bovine papillomavirus
CC type 4.
CC See also AAR42360.
CC
XX Sequence 98 AA:
SO

Query Match 100.0%; Score 92; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTLGIVCPICSQKP 17
   |||||
Db 82 limgtlgivcpicsqkp 98

RESULT 7
AAM46886
ID AAM46886 standard; Protein; 98 AA.
XX
AC AAM46886;
XX
DT 15-JUN-1998 (first entry)
XX
DE Amino acid sequence of the HPV-16 E7 oncoprotein.
XX
E7 oncoprotein; proliferative state; HPV; kinase activity;
KM cyclin/cyclin-dependent kinase; p21CIP1; interaction; inactivation;
KW cyclin/cyclin-dependent kinase inhibitor.
XX
OS Human papillomavirus.
XX
PN US5736318-A.
XX
PD 07-APR-1998.
XX
PF 17-MAR-1995; 95US-0406248.
XX
PR 17-MAR-1995; 95US-0406248.
XX
PA (HARD ) HARVARD COLLEGE.
XX (HARD ) UNIV HARVARD.
XX
PI Jones DL, Munger K;
XX
XX WPI; 1998-239202/21.
DR N-PSDB; AAV16717.
XX
XX Evaluation of proliferative state of cells transformed with human
PT papilloma virus - by determining cyclin-dependent kinase activity
PT induced by E7 onco-protein
XX
PS Disclosure: Columns 19-20; 14pp; English.
XX
XX The present sequence represents Human papillomavirus (HPV), strain 16,
CC E7 oncoprotein. The proliferative state of a cell transformed with
CC HPV can be evaluated in the following manner. Cyclin/cyclin-dependent
CC kinase complexes containing protein p21CIP1 (AAM46887-88) are isolated
CC from the transformed cell, and the HPV E7 oncoprotein added to the
CC isolated protein. Cyclin/cyclin-dependent kinase complexes are isolated

```

```

CC from an untransformed cell that is substantially homegenic with the
CC transformed cell, and the HPV E7 oncoprotein added. The kinase
CC activities of the 2 samples are measured, where a proliferating
CC transformed cell has a greater kinase activity than the untransformed
CC cell. The method is used for determining the extent of interaction and/or
CC inactivation between a cyclin/cyclin-dependent kinase inhibitor and the
CC HPV E7 oncoprotein and thus evaluating the proliferative state of a
CC transformed cell.
XX
XX Sequence 98 AA:
SO

Query Match 100.0%; Score 92; DB 19; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTLGIVCPICSQKP 17
   |||||
Db 82 limgtlgivcpicsqkp 98

RESULT 8
AAY08020
ID AAY08020 standard; Protein; 98 AA.
XX
AC AAY08020;
XX
DT 08-JUL-1999 (first entry)
XX
DE Human papilloma virus E7 protein.
XX
L1 protein; capsomer; virus; vaccine; infection; treatment; prevention;
KM cervical carcinoma; fusion protein; anti-capsid; antibody; E7 protein;
KM antigenicity.
XX
OS Human papilloma virus.
XX
PN W05918220-A1.
XX
PD 15-APR-1999.
XX
PF 06-OCT-1998; 98WO-US20965.
XX
PR 06-OCT-1997; 97US-0944368.
XX
PA (LOYO ) UNIV LOYOLA CHICAGO.
XX
PI Glesmann L, Mueller M;
XX
XX WPI; 1999-264026/22.
DR N-PSDB; AAX37567.
XX
XX Human papilloma virus (HPV) L1 fusion protein capsomers, used in
PT vaccines against HPV infection
XX
PS Disclosure: Page 33; 48pp; English.
XX
XX This invention describes novel vaccines comprising a human papilloma
CC virus (HPV) capsomer, composed of a HPV L1 fusion protein or a HPV L1
CC truncated protein. The vaccines can prevent or treat human papilloma
CC virus infection. Such therapeutic vaccinations can be used for relief
CC against, e.g. cervical carcinoma. Construction of chimeric proteins
CC comprising amino acid residues from L1 protein, and e.g. E6 or E7
CC protein, which give rise to chimeric capsomers, combines prophylactic and
CC therapeutic functions of a vaccine. Capsomers can promote elimination of
CC persistently infected cells. Capsomers can also escape neutralization
CC by pre-existing anti-capsid antibodies and hence possess longer
CC circulating half-life as compared to chimeric virus-like particles.
CC The fusion protein, which forms the capsomer, provides increased
CC antigenicity.
XX
XX Sequence 98 AA:
SO

```



Query Match 100.0%; Score 92; DB 20; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMGTLGIWPCISQKP 17  
 |||||  
 DB 82 limgtlgicpicsqkp 98

## RESULT 9

AAV58474  
 ID AAV58474 standard; protein; 98 AA.

AC AAV58474;

DT 10-APR-2000 (first entry)

DE Human papillomavirus (HPV) E7 oncoprotein.

DE HPV E7 oncoprotein; proteasome activity; degradation; virus component;

KW viral infection; inflammatory disease; anti-inflammatory; anti-HIV;

KW virucide.

OS Human papillomavirus.

PN W09966065-A2.

PD 23-DEC-1999.

PE 10-JUN-1999; 99WO-GB01840.

PF 13-JUN-1998; 98GB-0012756.

PR 13-JUN-1998; 98GB-0012757.

PR 13-JUN-1998; 98GB-0012758.

PR 13-JUN-1998; 98GB-0012759.

PR 13-JUN-1998; 98GB-0012760.

PA (BRID-) BRIDGEHEAD TECHNOLOGIES LTD.

PI Schmidt HP, Pettit F, Kioetzel PM, Jarrousse AS, Gaultier K;

PI Badouli S, Mouzeyar S, Nicolas P;

DR WPI: 2000-106109/09.

PT Novel assay methods for identifying compounds which modulate and/or

PT regulate proteasomal activity -

PS Disclosure: Page 8; 35pp; English.

XX The invention relates to a novel assay for identifying compounds which

CC inhibit viral replication and pathogenesis by increasing proteasomal

CC degradation of viral components or molecules induced by viral infection.

CC The method comprises reacting a compound with proteasomal protein (20S

CC proteasomes separately with or without 19S and 11S complexes), viral

CC gene product, and protein or peptide substrates; measuring protease

CC activity; and identifying the compound as an inhibitor if the protease

CC activity of the assay system is increased. The assays of the invention

CC can be used to identify compounds which inhibit viral replication and

CC pathogenesis, and modulate and/or regulate proteasome activity.

CC Compounds which increase proteasomal degradation of viral components or

CC molecules induced by viral infection are of value in the treatment of

CC viral disease. Compounds which modulate proteasomal nuclease activity

CC have use in the treatment of inflammatory disease, and AIDS in HIV

CC infected patients. The methods may also be used to generate resistance

CC to bacterial or viral damage. Sequences AAV58472-Y58474 represent

CC examples of viral proteins which affect proteasomal function.

XX Sequence 98 AA;

SO

Query Match 100.0%; Score 92; DB 21; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMGTLGIWPCISQKP 17  
 |||||  
 DB 82 limgtlgicpicsqkp 98

## RESULT 10

AAV57721  
 ID AAV57721 standard; protein; 98 AA.

AC AAV57721;

DT 14-MAR-2000 (first entry)

DE Human papillomavirus 16 E7 protein SEQ ID NO:4.

DE Human papillomavirus; HPV; L1 fusion protein; vaccine; cytostatic;

KW viral capsomere; virucide; dermatological; malignant tumour formation;

KW cervical cancer; cervical intraepithelial neoplasia; genital wart;

OS Human papillomavirus.

PN CA2229955-A1.

PD 20-AUG-1999.

PE 20-FEB-1998; 98CA-2229955.

PR 20-FEB-1998; 98CA-2229955.

PR (MED-) MEDIGENE GMBH.

PI Burger A, Hallek M;

PI WPI: 2000-063092/06.

DR N-PSDB; AA248175.

PT Fusion proteins comprising papillomavirus specific proteins useful for

PT vaccinating against malignant tumors of the anogenital tract such as

PS cervical carcinomas -

PS Example 1; Page 34; 46pp; English.

XX The present invention describes a fusion protein comprising 2 amino

CC acids sequences from 2 different papillomavirus specific (PVS) proteins.

CC The fusion protein may be administered for preventing and treating

CC papillomavirus infections in humans and animals. Papillomaviruses are

CC implicated in the pathology of malignant tumour formation in the

CC anogenital tract (of these tumours, cervical cancer is the most frequent

CC (50000 cases/year) and in the formation of precursor lesions of

CC cervical intraepithelial neoplasia (CIN). Papillomaviruses also cause

CC benign genital warts such as condylomata acuminata. However, the type

CC and severity of disease caused by the papillomavirus is dependent on the

CC strain causing the infection. The present sequence represents the human

CC papillomavirus 16 E7 protein.

XX Sequence 98 AA;

SO

Query Match 100.0%; Score 92; DB 21; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMGTLGIWPCISQKP 17  
 |||||  
 DB 82 limgtlgicpicsqkp 98

RESULT 11  
 AAV50703  
 ID AAV50703 standard; protein; 98 AA.

```

XX AAY50703;
XX 04-FEB-2000 (first entry)
XX
XX HPV16 E7 protein.
XX
XX E7 protein; Immunogenic; active immunization.
XX
XX Human papillomavirus.
XX
XX MO9955876-A2.
XX
XX 04-NOV-1999.
XX
XX 30-APR-1999; 99WO-DE01331.
XX
XX 30-APR-1998; 98DE-1019476.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Glassmann L, Jochmus I;
XX
XX WPI: 2000-023362/02.
XX
XX N-PSDB; AA224129.
XX
XX Immunogenic protein with altered biological function, useful for active
XX immunization
XX
XX Example 1; Page 25; 33pp; German.
XX
XX This invention describes the construction of a novel polypeptide with
XX immunogenic and altered biological function of a protein, where the
XX polypeptide has regions of the protein of about 10-40 amino acids in a
XX different order. The polypeptide, or its DNA, is useful for active
XX immunization without cross-reactivity and problems associated with the
XX biological function of the protein. This sequence represents the human
XX papillomavirus type 16 E7 protein which is used to illustrate the method
XX of the invention.
XX
XX Sequence 98 AA:
XX

```

```

Query Match 100.0%; Score 92; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LLMGTLGIVPCISQKP 17
Db 82 LLMGTLGIVPCISQKP 98

```

```

RESULT 12
AAB86332
ID AAB86332 standard; Protein: 98 AA.
XX
XX AAB86332;
XX
XX 18-SEP-2001 (first entry)
XX
XX HPV 16 E7 peptide fragment.
XX
XX Fusion protein; VP2; E7; cell import signal; cell export signal;
XX antigen; immunization; infection-induced auto-immune disease;
XX tumor disease.
XX
XX Human papillomavirus.
XX
XX WO200151516-A2.
XX
XX 19-JUL-2001.
XX
XX 15-JAN-2001; 2001WO-DE00134.
XX

```

```

XX 13-JAN-2000; 2000DE-1001230.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Mueller M, Michel N, Osen W, Gissmann L, Zentgraf H;
XX
XX WPI: 2001-442135/47.
XX
XX Identifying an immunization agent comprising cell import and/or
XX export signal sequences and an antigen for immunizing against
XX infection-induced auto-immune and tumor disease
XX
XX Disclosure: Fig 4; 23pp; German.
XX
XX This invention describes a fusion protein comprising cell import and/or
XX export signal sequences and an antigen which is suitable for immunizing
XX an individual against a disease, together with a DNA that codes for said
XX protein. The invention also relates to the use of the protein (II) and
XX its encoding DNA (I) for immunizing an individual against diseases. In
XX particular against infection-induced auto-immune and tumor disease. This
XX sequence represents a fragment of the human papillomavirus E7 protein
XX used in the preparation of the fusion constructs described in the method
XX of the invention.
XX
XX Sequence 98 AA:
XX

```

```

Query Match 100.0%; Score 92; DB 22; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LLMGTLGIVPCISQKP 17
Db 82 LLMGTLGIVPCISQKP 98

```

```

RESULT 13
AAB98421
ID AAB98421 standard; Protein: 98 AA.
XX
XX AAB98421;
XX
XX 22-AUG-2001 (first entry)
XX
XX Human papillomavirus protein HPV16 E7.
XX
XX Human papillomavirus; human leukocyte antigen; HLA; immune response;
XX HPV; epitope; T cell; identification; vaccine; infection; genital wart;
XX neoplastic growth; antiviral.
XX
XX Human papillomavirus.
XX
XX WO200141799-A1.
XX
XX 14-JUN-2001.
XX
XX 11-DEC-2000; 2000WO-US33549.
XX
XX 10-DEC-1999; 99US-0172705.
XX
XX 15-AUG-2000; 2000US-0641528.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Settle A, Sidney J, Southwood S, Chesnut R, Cells E, Grey HM;
XX
XX WPI: 2001-381497/40.
XX
XX An isolated human papilloma virus (HPV) epitope, useful in vaccines for
XX treating HPV infections -
XX
XX Disclosure: Page 21; 756pp; English.
XX

```

CC The present invention describes an isolated prepared human papillomavirus  
 CC (HPV) epitope (1). (1) has antiviral activity, and can be used in  
 CC vaccine production. Peptides and corresponding nucleic acid compositions  
 CC from the present invention are useful for stimulating an immune response  
 CC to HPV by stimulating the production of CTL or HTL responses,  
 CC specifically in the treatment or prophylaxis of HPV infection, in persons  
 CC who have not manifested symptoms e.g. genital warts or neoplastic growth.  
 CC The peptides can also be used in a tetramer staining assay to assess  
 CC peripheral blood mononuclear cells for the presence of antigen-specific  
 CC cells following exposure to a pathogen or immunogen, and as reagents to  
 CC evaluate immune recall responses or evaluate the efficacy of a vaccine.  
 CC The vaccine compositions are useful for removing warts or treating HPV  
 CC infections. The epitopes for inclusion in an epitope-base vaccine may  
 CC be selected from conserved regions of viral or tumour-associated  
 CC antigens, which reduces the likelihood of escape mutants, also  
 CC immunosuppressive epitopes that may be present in whole antigens can be  
 CC avoided with the use of epitope-base vaccines. An additional advantage  
 CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
 CC the composition of the epitopes achieving enhanced immunogenicity, the  
 CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
 CC to AAB98477 represent polypeptide sequences used in the exemplification  
 CC of the present invention.

CC Sequence 98 AA;  
 XX

Query Match 100.0%; Score 92; DB 22; Length 98;  
 Best Local Similarity 100.0%; Pred. NO. 1.7e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTIGIVCPICSQKP 17  
 |||||  
 DB 82 Llmgtlgtvcpicsqkp 98

RESULT 14

ID AAU01718 standard; Protein; 98 AA.

XX AAU01718;

XX 29-AUG-2001 (first entry)

XX Human papilloma virus-16 (HPV-16) E7 antigen.

XX Mouse; granulocyte-macrophage-colony stimulating factor; GM-CSF;

XX chimeric; heat shock protein; HSP; P1t-3 ligand; FL; exotoxin A;

XX E7A d1t; antigenic; immunogenic; cytotoxic T cell response; tumour;

XX vaccine; immunotherapy; HPV-16; E7 antigen.

XX Human papillomavirus.

XX WO200129233-A2.

XX 26-APR-2001.

XX 20-OCT-2000; 2000WO-USA1422.

XX 20-OCT-1999; 99US-0421608.

XX 09-FEB-2000; 2000US-0501097.

XX (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

XX Wu T, Hung C;

XX WPI; 2001-290921/30.

XX N-PSDB; AAS02608.

XX New chimeric polypeptide, useful as anti-tumour vaccines, comprises  
 PT carboxy terminal fragment of heat shock protein, P1t-3 ligand or  
 PT cytoplasmic translocation domain of Pseudomonas exotoxin A and  
 PT antigenic polypeptide.

P5 Example 1; Page 39; 110pp; English.

XX The sequence represents the amino acid sequence of human papilloma  
 CC virus-16 (HPV-16) E7 antigen used in construction of a chimeric  
 CC polypeptide comprising: (a) a first polypeptide domain containing a  
 CC carboxy terminal fragment of a heat shock protein (HSP), an  
 CC P1t-3 ligand (FL), a cytoplasmic translocation domain of a Pseudomonas  
 CC exotoxin A (E7A d1t), or a granulocyte-macrophage colony stimulating  
 CC factor (GM-CSF); and (b) a second polypeptide domain containing an  
 CC antigenic polypeptide. A composition comprising the chimeric  
 CC polypeptide is useful for inducing an immune response such as a cytotoxic  
 CC T cell response. The nucleic acid or vector encoding the chimeric  
 CC polypeptide present in the composition is administered as naked DNA by  
 CC gene gun or equivalent, or by liposomal formulation. These are thus  
 CC useful for vaccinating a mammal against infection by inducing an immune  
 CC response to a pathogen. Preferably they are useful for vaccinating a  
 CC mammal against a tumour antigen. The compositions and methods are useful  
 CC for stimulating or enhancing the immunogenicity of a selected antigen or  
 CC stimulating or enhancing a cellular immune response specific for that  
 CC antigen. The chimeric nucleic acid molecules and vaccination methods,  
 CC yield potent antigen-specific immunotherapy. The polynucleotides and DNA  
 CC vaccines can induce a cellular immune response that is at least 40 fold  
 CC more potent than conventional DNA vaccines. The vaccines are safe and  
 CC useful for administration to domesticated or agricultural animals, as  
 CC well as humans, and have low immunogenicity.

XX Sequence 98 AA;  
 XX

Query Match 100.0%; Score 92; DB 22; Length 98;  
 Best Local Similarity 100.0%; Pred. NO. 1.7e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTIGIVCPICSQKP 17  
 |||||  
 DB 82 Llmgtlgtvcpicsqkp 98

RESULT 15

ID AAB67546 standard; peptide; 98 AA.

XX AAB67546;

XX 29-MAY-2001 (first entry)

XX Amino acid sequence of a E7 fragment that activates CDK2.

XX E7-induced CDK2 kinase; CDK2 substrate; papillomavirus; E7;

XX viral infection.

XX Papillomavirus sylv1ag1.

XX WO200114584-A2.

XX 01-MAR-2001.

XX 25-AUG-2000; 2000WO-US23487.

XX 25-AUG-1999; 99US-0382616.

XX (PHAA ) PHARMACIA & UPJOHN.

XX Fisher C, He W;

XX WPI; 2001-226622/23.

XX Identifying inhibitors of human papilloma virus E7 protein-induced  
 PT increase in CDK2 kinase activity for use as antiviral agent, by  
 PT measuring kinase activity on CDK2 substrate in presence of E7 protein,  
 PT test compound -

XX Claim 2; Page 40; 49pp; English.

XX The specification describes a method for identifying inhibitors of  
CC E7-induced CDK2 kinase activity. The method comprises measuring CDK2  
CC kinase activity on a CDK2 substrate in presence of human papillomavirus  
CC (HPV) E7 (or fragment), in presence and absence of a test compound and  
CC measuring phosphorylation of substrate. The method is useful for  
CC identifying an inhibitor of E7-induced CDK2 kinase activity. The  
CC identified inhibitor is useful for reducing or inhibiting HPV  
CC proliferation in an individual. Inhibitors of E7 binding to the CDK2  
CC complex are useful in preparing medicament for ameliorating viral  
CC infection e.g. HPV infection, adenoviral or SV40 infection. The present  
CC sequence represents a papillomavirus E7 fragment, which is used in the  
CC method of the invention.

XX  
SQ Sequence 98 AA;

Query Match 100.0%; Score 92; DB 22; Length 98;  
Best Local Similarity 100.0%; Pred. No. 1.7e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMCTGIVCPICSQKP 17  
|||||  
Db 82 llmgtlgivcpicsqkp 98

Search completed: February 13, 2002, 09:59:18  
Job time: 332 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 09:59:57 ; Search time 32.4 seconds  
(without alignments)  
11.807 Million cell updates/sec

Title: US-09-664-225-69

Perfect score: 92  
Sequence: 1 LMGTLGIVCICSKP 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents\_AA:\*

1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	21	2	US-08-934-915-50 Sequence 50, Appl
2	92	100.0	21	2	US-08-934-915-157 Sequence 157, App
3	92	100.0	30	2	US-08-934-915-54 Sequence 54, Appl
4	92	100.0	98	1	US-08-406-248-6 Sequence 6, Appl1
5	92	100.0	98	4	US-08-075-541D-42 Sequence 42, Appl
6	92	100.0	98	4	US-09-382-616A-1 Sequence 1, Appl1
7	92	100.0	98	4	US-08-944-368A-4 Sequence 4, Appl1
8	92	100.0	172	3	US-08-860-165-14 Sequence 14, Appl
9	92	100.0	172	4	US-09-359-382-14 Sequence 10, Appl
10	92	100.0	253	2	US-08-459-818-20 Sequence 20, Appl
11	92	100.0	253	2	US-08-889-666-20 Sequence 20, Appl
12	92	100.0	253	2	US-08-465-078-20 Sequence 20, Appl
13	92	100.0	253	2	US-08-725-776-20 Sequence 20, Appl
14	92	100.0	253	2	US-08-488-062-20 Sequence 20, Appl
15	92	100.0	263	1	US-08-117-083-9 Sequence 9, Appl
16	92	100.0	266	3	US-08-860-165-10 Sequence 10, Appl
17	92	100.0	266	4	US-09-359-382-10 Sequence 10, Appl
18	85	92.4	26	4	US-09-169-425C-25 Sequence 25, Appl
19	85	92.4	26	4	US-08-075-541D-40 Sequence 40, Appl
20	81	88.0	15	3	US-08-159-339A-1168 Sequence 1168, Ap
21	77	83.7	14	4	US-09-169-425C-32 Sequence 32, Appl
22	72	78.3	13	3	US-08-159-339A-1167 Sequence 1167, Ap
23	71	77.2	13	3	US-08-948-378A-3 Sequence 3, Appl1
24	71	77.2	13	3	US-09-169-425C-3 Sequence 3, Appl1
25	71	77.2	20	4	US-08-075-541D-50 Sequence 50, Appl
26	67	72.8	12	3	US-08-948-378A-16 Sequence 16, Appl
27	67	72.8	12	4	US-09-169-425C-16 Sequence 16, Appl

28	67	72.8	13	3	US-08-948-378A-4 Sequence 4, Appl1
29	67	72.8	13	3	US-08-948-378A-19 Sequence 19, Appl
30	67	72.8	13	4	US-09-169-425C-4 Sequence 4, Appl1
31	67	72.8	13	4	US-09-169-425C-19 Sequence 19, Appl1
32	67	72.8	38	3	US-08-948-378A-6 Sequence 6, Appl1
33	67	72.8	38	4	US-09-169-425C-6 Sequence 6, Appl1
34	67	72.8	601	2	US-08-606-288-10 Sequence 10, Appl
35	67	72.8	601	2	US-08-606-288-10 Sequence 10, Appl
36	67	72.8	601	3	US-09-347-483-7 Sequence 7, Appl1
37	67	72.8	601	3	US-09-347-483-10 Sequence 10, Appl
38	63	68.5	11	4	US-09-169-425C-33 Sequence 33, Appl
39	58	63.0	11	4	US-09-169-425C-31 Sequence 31, Appl
40	57	62.0	10	3	US-08-159-339A-572 Sequence 572, App
41	52	56.5	9	3	US-08-948-378A-2 Sequence 2, Appl1
42	52	56.5	9	4	US-09-169-425C-2 Sequence 2, Appl1
43	52	56.5	9	5	PCT-US95-02121-68 Sequence 68, Appl
44	51	55.4	9	3	US-08-159-339A-83 Sequence 83, Appl
45	51	55.4	9	4	US-09-169-425C-22 Sequence 22, Appl

## ALIGNMENTS

RESULT 1  
US-08-934-915-50  
; Sequence 50, Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; APPLICANT: DILLNER, LENA  
; APPLICANT: CHENG, HWEI-MING  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
; TITLE OF INVENTION: USEFUL IN IMMUNOSSAY FOR  
; NUMBER OF SEQUENCES: 193  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: MASON & ASSOCIATES, P.A.  
; STREET: 17757 Q.S. HWY. 19 NORTH, SUITE 500  
; CITY: CLEARWATER  
; STATE: FLORIDA  
; COUNTRY: U.S.A.  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 3.0  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/934,915  
; FILING DATE: 22-SEP-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/949,836  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LOUISE A. FOUTCH  
; REGISTRATION NUMBER: 37,133  
; REFERENCE/DOCKET NUMBER: 1946.6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 813-538-3800  
; TELEFAX: 813-538-3820  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-934-915-50  
Query Match 100.0%; Score 92; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLMGTLGIVCPICSKP 17  
|||||  
Db 5 LLMGTLGIVCPICSKP 21

RESULT 2  
US-08-934-915-157  
Sequence 157; Application US/08934915  
Patent No. 5932412  
GENERAL INFORMATION:  
APPLICANT: DILLNER, JOAKIM  
APPLICANT: DILLNER, LENA  
APPLICANT: CHENG, HWE-MING  
TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
NUMBER OF SEQUENCES: 193  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MASON & ASSOCIATES, P.A.  
STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
CITY: CLEARWATER  
STATE: FLORIDA  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 3.0  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/934,915  
FILING DATE: 22-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/949,836  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LOUISE A. FOUTCH  
REGISTRATION NUMBER: 37,133  
REFERENCE/DOCKET NUMBER: 1946.6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 813-538-3800  
TELEFAX: 813-538-3820  
INFORMATION FOR SEQ. ID NO.: 157:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-934-915-157

Query Match 100.0%; Score 92; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLMGTLGIVCPICSKP 17  
|||||  
Db 5 LLMGTLGIVCPICSKP 21  
RESULT 3  
US-08-934-915-54  
Sequence 54; Application US/08934915  
Patent No. 5932412  
GENERAL INFORMATION:  
APPLICANT: DILLNER, JOAKIM  
APPLICANT: DILLNER, LENA

APPLICANT: CHENG, HWE-MING  
TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
NUMBER OF SEQUENCES: 193  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MASON & ASSOCIATES, P.A.  
STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
CITY: CLEARWATER  
STATE: FLORIDA  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 3.0  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/934,915  
FILING DATE: 22-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/949,836  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LOUISE A. FOUTCH  
REGISTRATION NUMBER: 37,133  
REFERENCE/DOCKET NUMBER: 1946.6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 813-538-3800  
TELEFAX: 813-538-3820  
INFORMATION FOR SEQ. ID NO.: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-934-915-54

Query Match 100.0%; Score 92; DB 2; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2.5e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLMGTLGIVCPICSKP 17  
|||||  
Db 14 LLMGTLGIVCPICSKP 30

RESULT 4  
US-08-406-248-6  
Sequence 6; Application US/08406248  
Patent No. 5736318  
GENERAL INFORMATION:  
APPLICANT: Munger, Karl  
APPLICANT: Jones, D. Leanne  
TITLE OF INVENTION: METHOD AND KIT FOR EVALUATING  
TITLE OF INVENTION: TRANSFORMED CELLS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ann-Louise Kerner, Ph.D., Lappin & Kusner  
STREET: 200 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/406,248  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: McDaniel, Patricia A.  
REGISTRATION NUMBER: 33,194  
REFERENCE/DOCKET NUMBER: HAZ-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-330-1300  
TELEFAX: 617-330-1311  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 98 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-406-248-6

Query Match 100.0%; Score 92; DB 1; Length 98;  
Best Local Similarity 100.0%; Pred. No. 8.8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMGTGIVCPICSKP 17  
|||||  
Db 82 LMGTGIVCPICSKP 98

RESULT 5  
US-08-075-541D-42  
Sequence 42, Application US/08075541D

Patent No. 6183745

GENERAL INFORMATION:

APPLICANT: TINDLE, ROBERT

APPLICANT: FERNANDO, GERMAIN

APPLICANT: FRAZER, IAN

TITLE OF INVENTION: SUBUNIT PAPILLOMA VIRUS VACCINE AND

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSEE: PANITCH SCHWARZ JACOBS & NADEL, P.C.

STREET: 1601 MARKET STREET, 36TH FLOOR

CITY: PHILADELPHIA

STATE: PENNSYLVANIA

COUNTRY: USA

ZIP: 19103-2398

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/075,541D

FILING DATE: 10-JUN-1993

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PK 3876

FILING DATE: 12-DEC-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/AU91/00575

FILING DATE: 12-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: NADEL, ALAN S

REGISTRATION NUMBER: 27,363

REFERENCE/DOCKET NUMBER: 8795-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-567-2020

TELEFAX: 215-567-2991

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:

LENGTH: 98 amino acids

TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-075-541D-42

Query Match 100.0%; Score 92; DB 4; Length 98;  
Best Local Similarity 100.0%; Pred. No. 8.8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMGTGIVCPICSKP 17  
|||||  
Db 82 LMGTGIVCPICSKP 98

RESULT 6  
US-09-382-616A-1  
Sequence 1, Application US/09382616A

Patent No. 6200746

GENERAL INFORMATION:

APPLICANT: Fisher, Christopher

APPLICANT: He, Wankia

TITLE OF INVENTION: Methods to Identify Anti-Viral Agents

FILE REFERENCE: 28341/6216

CURRENT APPLICATION NUMBER: US/09/382,616A

CURRENT FILING DATE: 1999-08-25

PRIOR APPLICATION NUMBER: 09/382,616

PRIOR FILING DATE: 1999-08-25

NUMBER OF SEQ ID NOS: 43

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 1

LENGTH: 98

TYPE: PRT

ORGANISM: Papillomavirus sylvilagi

US-09-382-616A-1

Query Match 100.0%; Score 92; DB 4; Length 98;  
Best Local Similarity 100.0%; Pred. No. 8.8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMGTGIVCPICSKP 17  
|||||  
Db 82 LMGTGIVCPICSKP 98

RESULT 7  
US-08-944-368A-4  
Sequence 4, Application US/08944368A

Patent No. 6228368

GENERAL INFORMATION:

APPLICANT: Gissman, et al.

TITLE OF INVENTION: Papilloma Virus Capsomere Vaccine

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

STREET: 233 South Wacker Drive, 6300 Sears Tower

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/944,368A

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

```

: NAME: Williams Jr., Joseph A.
: REGISTRATION NUMBER: 38,659
: REFERENCE/DOCKET NUMBER: 27013/34028
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 312-474-6300
:   TELEFAX: 312-474-0448
: INFORMATION FOR SEQ ID NO: 4:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 98 amino acids
:     TYPE: amino acid
:     TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-944-368A-4

Query Match          100.0%; Score 92; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 8.8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTGIVCPICSQKP 17
Db 82 LLMGTGIVCPICSQKP 98

RESULT 8
US-08-860-165-14
: Sequence 14, Application US/08860165A
: Patent No. 6004557
: GENERAL INFORMATION:
:   APPLICANT: EDWARDS, Stirling John
:   APPLICANT: COX, John Cooper
:   APPLICANT: WEBB, Elizabeth Ann
:   TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
:   FILE REFERENCE: 17227/130
:   CURRENT APPLICATION NUMBER: US/08/860,165A
:   EARLIER FILING DATE: 1997-09-22
:   EARLIER APPLICATION NUMBER: PCT/AU95/00868
:   EARLIER FILING DATE: 1995-12-20
:   EARLIER APPLICATION NUMBER: AU PN0157
:   NUMBER OF SEQ ID NOS: 15
:   SOFTWARE: PatentIn Ver. 2.0
:   SEQ ID NO 14
:   LENGTH: 172
:   TYPE: PRT
:   ORGANISM: Artificial Sequence
:   FEATURE:
:   OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
: US-08-860-165-14

Query Match          100.0%; Score 92; DB 3; Length 172;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTGIVCPICSQKP 17
Db 50 LLMGTGIVCPICSQKP 66

RESULT 9
US-09-359-382-14
: Sequence 14, Application US/09359382
: Patent No. 6306397
: GENERAL INFORMATION:
:   APPLICANT: EDWARDS, Stirling John
:   APPLICANT: COX, John Cooper
:   APPLICANT: WEBB, Elizabeth Ann
:   APPLICANT: FRAZER, Ian
:   TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
:   FILE REFERENCE: 017227/0148
:   CURRENT APPLICATION NUMBER: US/09/359,382
```

```

: CURRENT FILING DATE: 1999-07-23
: EARLIER APPLICATION NUMBER: US 08/860,165
: EARLIER FILING DATE: 1997-09-22
: EARLIER APPLICATION NUMBER: PCT/AU95/00868
: EARLIER FILING DATE: 1995-12-20
: EARLIER APPLICATION NUMBER: AU PN0157/94
: EARLIER FILING DATE: 1994-12-20
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.0
:   SEQ ID NO 14
:   LENGTH: 172
:   TYPE: PRT
:   ORGANISM: Human papillomavirus type 16
: US-09-359-382-14

Query Match          100.0%; Score 92; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTGIVCPICSQKP 17
Db 50 LLMGTGIVCPICSQKP 66

RESULT 10
US-08-459-818-20
: Sequence 20, Application US/08459818
: Patent No. 5851795
: GENERAL INFORMATION:
:   APPLICANT: Linsley, Peter S.
:   APPLICANT: Ledbetter, Jeffrey A.
:   APPLICANT: Damle, Nitin K.
:   APPLICANT: Brady, William
:   TITLE OF INVENTION: CTLA4 Receptor and uses thereof
:   NUMBER OF SEQUENCES: 27
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: Merchant & Gould
:     STREET: 11150 Santa Monica Blvd., Suite 400
:     CITY: Los Angeles
:     STATE: California
:     COUNTRY: USA
:     ZIP: 90025
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: FastSeq 2.0
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/08/459,818
:     FILING DATE: 02-JUN-1995
:     CLASSIFICATION: 435
:     ATTORNEY/AGENT INFORMATION:
:       NAME: Adriano, Sarah B.
:       REGISTRATION NUMBER: 34,470
:       REFERENCE/DOCKET NUMBER: 30436.350S02
:       TELECOMMUNICATION INFORMATION:
:         TELEPHONE: 310-445-9031
:         TELEFAX: 310-445-9031
:       INFORMATION FOR SEQ ID NO: 20:
:         SEQUENCE CHARACTERISTICS:
:           LENGTH: 253 amino acids
:           TYPE: amino acid
:           STRANDEDNESS:
:           TOPOLOGY: linear
:         MOLECULE TYPE: protein
: US-08-459-818-20

Query Match          100.0%; Score 92; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



QY 1 LLMGTGIVCPSQKP 17  
|||||  
DB 237 LLMGTGIVCPSQKP 253

RESULT 11  
US-08-889-666-20

; Sequence 20, Application US/08889666  
; Patent No. 588579

; GENERAL INFORMATION:

; APPLICANT: Linsley, Peter S.  
; APPLICANT: Ledbetter, Jeffrey A.

; APPLICANT: Damle, Nitin K.  
; APPLICANT: Brady, William

; APPLICANT: Kiener, Peter A.  
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof

; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Merchant & Gould  
; STREET: 1150 Santa Monica Blvd., Suite 400

; CITY: Los Angeles  
; STATE: California

; COUNTRY: USA  
; ZIP: 90025

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/889,666

; FILING DATE: 08-JUL-1997  
; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/375390

; FILING DATE: 18-JAN-1995  
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: Adriano, Sarah B.

; REGISTRATION NUMBER: 34,470  
; REFERENCE/DOCKET NUMBER: 30436-35US01

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 310-445-9031

; TELEFAX: 310-445-1140  
; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids

; TYPE: amino acid  
; STRANDEDNESS:

; TOPOLOGY: linear  
; MOLECULE TYPE: protein

; US-08-889-666-20

Query Match 100.0%; Score 92; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTGIVCPSQKP 17  
|||||

DB 237 LLMGTGIVCPSQKP 253

RESULT 12  
US-08-465-078-20

; Sequence 20, Application US/08465078  
; Patent No. 5885796

; GENERAL INFORMATION:

; APPLICANT: Linsley, Peter S.  
; APPLICANT: Ledbetter, Jeffrey A.

; APPLICANT: Damle, Nitin K.  
; APPLICANT: Brady, William

; APPLICANT: Kiener, Peter A.

; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof

; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Merchant & Gould  
; STREET: 1150 Santa Monica Blvd., Suite 400

; CITY: Los Angeles  
; STATE: California

; COUNTRY: USA  
; ZIP: 90025

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,078

; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/375390

; FILING DATE: 18-JAN-1995  
; ATTORNEY/AGENT INFORMATION:

; NAME: Adriano, Sarah B.  
; REGISTRATION NUMBER: 34,470

; REFERENCE/DOCKET NUMBER: 30436-35US01  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 310-445-9031  
; TELEFAX: 310-445-1140

; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 253 amino acids  
; TYPE: amino acid

; STRANDEDNESS:  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
; US-08-465-078-20

Query Match 100.0%; Score 92; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTGIVCPSQKP 17  
|||||

DB 237 LLMGTGIVCPSQKP 253

RESULT 13  
US-08-725-776-20

; Sequence 20, Application US/08725776  
; Patent No. 5968510

; GENERAL INFORMATION:

; APPLICANT: Linsley, Peter S.  
; APPLICANT: Ledbetter, Jeffrey A.

; APPLICANT: Damle, Nitin K.  
; APPLICANT: Brady, William

; APPLICANT: Kiener, Peter A.  
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof

; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Merchant & Gould  
; STREET: 1150 Santa Monica Blvd., Suite 400

; CITY: Los Angeles  
; STATE: California

; COUNTRY: USA  
; ZIP: 90025

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/725,776

FILED DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375390  
FILING DATE: 18-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Adriano, Sarah B.  
REGISTRATION NUMBER: 34,470  
REFERENCE/DOCKET NUMBER: 30436-350S01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 310-445-1140  
TELEFAX: 310-445-9031  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-725-776-20

Query Match 100.0%; Score 92; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTGIVCPICSKP 17  
|||||  
Db 237 LLMGTGIVCPICSKP 253

RESULT 14  
US-08-488-062-20  
Sequence 20, Application US/08488062  
Patent No. 5977318  
GENERAL INFORMATION:  
APPLICANT: Linsley, Peter S.  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Dame, Nalin K.  
APPLICANT: Brady, William  
APPLICANT: Kienner, Peter A.  
TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 11150 Santa Monica Blvd., Suite 400  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,062  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375390  
FILING DATE: 18-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Adriano, Sarah B.  
REGISTRATION NUMBER: 34,470  
REFERENCE/DOCKET NUMBER: 30436-350S01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 310-445-1140  
TELEFAX: 310-445-9031  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-062-20

Query Match 100.0%; Score 92; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTGIVCPICSKP 17  
|||||  
Db 237 LLMGTGIVCPICSKP 253

RESULT 15  
US-08-117-083-9  
Sequence 9, Application US/08117083  
Patent No. 5719054  
GENERAL INFORMATION:  
APPLICANT: Boursnell, Michael E.  
APPLICANT: Inglis, Stephen C.  
APPLICANT: Munro, Alan J.  
TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human  
NUMBER OF SEQUENCES: 70  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Walter H. Dreger  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/117,083  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Walter H.  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-58783  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..263  
OTHER INFORMATION: /note- "Xaa refers to stop codon in  
OTHER INFORMATION: the open reading frame."  
US-08-117-083-9

Query Match 100.0%; Score 92; DB 1; Length 263;  
Best Local Similarity 100.0%; Pred. No. 2.5e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMGTGIVCPICSKP 17  
|||||  
Db 243 LLMGTGIVCPICSKP 259

Search completed: February 13, 2002, 09:59:57  
Job time: 346 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 10:00:43 : Search time 39.34 Seconds  
(without alignments)  
32.917 Million cell updates/sec

Title: US-09-664-225-69

Perfect score: 92

Sequence: 1 LMGTGIVCPICSQKP 17

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	98	1 W7MLH1	E7 protein - human
2	74	80.4	98	1 W7MLI1	E7 protein - human
3	70	76.1	99	1 W7MLJ5	E7 protein - human
4	67	72.8	98	1 W7MLI6	E7 protein - human
5	65	70.7	101	1 W7MLI3	E7 protein - human
6	64	69.6	55	2 S19907	E7-C protein - hum
7	64	69.6	97	1 W7MLI3	E7 protein - human
8	61	66.3	98	1 W7MLJ1	E7 protein - human
9	61	66.3	113	1 W7MLR1	E7 protein - human
10	59	64.1	93	1 W7MLI2	E7 protein - human
11	59	64.1	98	1 W7MLI5	E7 protein - human
12	58	63.0	111	2 S36585	E7 protein - human
13	58	63.0	97	2 S36516	E7 protein - human
14	58	63.0	104	2 S36510	E7 protein - human
15	57	62.0	98	1 W7MLC1	E7 protein - human
16	57	62.0	111	2 S36556	E7 protein - human
17	56	60.9	105	2 S36528	E7 protein - human
18	54	58.7	105	2 S36504	E7 protein - human
19	52	56.5	105	2 B44890	E7 protein - human
20	52	56.5	105	2 S36580	E7 protein - human
21	52	56.5	105	2 S36580	E7 protein - human
22	51	55.4	86	2 S36533	Probable RING zinc
23	51	55.4	99	2 S36574	E7 protein - human
24	51	55.4	101	1 W7MLI1	E7 protein - human
25	49	53.3	93	1 W7ML	E7 protein - human
26	48	52.2	95	2 S36480	E7 protein - human
27	46	50.0	109	1 W7MLI3	E7 protein - human
28	45	48.9	995	2 T13648	mitosis initiation
29	45	48.9	708	2 A38436	mitosis initiation

30	45	48.9	1006	2 S74992	hypothetical prote
31	45	48.9	1496	1 CGHU2V	collagen alpha 2(V
32	44	47.8	125	2 C64777	probable membrane
33	44	47.8	125	2 E85544	probable gene 58 p
34	44	47.8	252	2 D85760	probable oxidoredu
35	44	47.8	365	2 B64875	probable dehydroge
36	44	47.8	365	2 C85933	membrane-bound 1Y
37	44	47.8	365	2 A65064	membrane-bound 1Y
38	44	47.8	449	2 B86763	hypothetical prote
39	44	47.8	639	2 T33166	hypothetical prote
40	43.5	47.3	93	2 A86491	CT001 hypothetical
41	43.5	47.3	93	2 F72130	conserved hypothet
42	43	46.7	92	1 S15622	E7 protein - human
43	43	46.7	93	2 S36591	E7 protein - human
44	43	46.7	105	1 W7MLI8	E7 protein - human
45	43	46.7	183	2 D72003	hypothetical prote

## ALIGNMENTS

RESULT 1  
W7MLH1  
E7 protein - human papillomavirus type 16  
C:Species: human papillomavirus type 16  
C>Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 20-Aug-1999  
C:Accession: A03688; S12367; T10428  
R:Seedorf, K.; Kramer, G.; Durst, M.; Suhai, S.; Roweckamp, W.G.  
Virology 145, 181-185, 1985  
A:Title: Human papillomavirus type 16 DNA sequence.  
A:Reference number: A22355; M0ID:85246220  
A:Accession: A03688  
A:Molecule type: DNA  
A:Residues: 1-98 <SEQ>  
A:Cross-references: GB:K02718; NID:9333031; PIDN:AAA6940.1; PID:9333033  
R:Barbosa, M.S.; Edmonds, C.; Fisher, C.; Schiller, J.T.; Lowy, D.R.; Vousden, K.H.  
EMBO J. 9, 153-160, 1990  
A:Title: The region of the HPV E7 oncoprotein homologous to adenovirus E1a and SV40 1  
A:Reference number: S12367; M0ID:90107938  
A:Accession: S12367  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-98 <BAR>  
R:Kennedy, I.M.; Haddow, J.K.; Clements, J.B.  
J. Virol. 65, 2093-2097, 1991  
A:Title: A negative element in the human poapillomavirus type 16 genome acts at the 1  
A:Reference number: Z17014; M0ID:91162763  
A:Accession: T10428  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-98 <KEN>  
A:Cross-references: EMBL:K02718; NID:9333031; PIDN:AAA6940.1; PID:9333033  
C:Gene: E7  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation; zinc finger  
F:58-94/Region: zinc finger CCCC motif

Query Match 100.0%; Score 92; DB 1; Length 98;  
Best local Similarity 100.0%; Pred. No. 7.2e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMGTGIVCPICSQKP 17  
|||||  
DB 82 LMGTGIVCPICSQKP 98

RESULT 2  
W7MLI1  
E7 protein - human papillomavirus type 11  
C:Species: human papillomavirus type 11  
C>Date: 13-Aug-1986 #sequence\_revision 13-Aug-1986 #text\_change 16-Jul-1999

C:Accession: A03690  
 C:Dartmann, K.; Schwarz, E.; Glasmann, L.; zur Hausen, H.  
 Virology 151, 124-130, 1986  
 A>Title: The nucleotide sequence and genome organization of human papilloma virus type 1  
 A:Reference number: A94338; MUID:86181601  
 A:Accession: A03690  
 A:Molecule type: DNA  
 A:Residues: 1-98 <DAS>  
 A:Cross-references: GB:M4119; NID:q333026; PTDN:AAA46928.1; PID:g496194  
 C:Superfamily: papillomavirus E7 protein  
 C:Keywords: DNA binding; early protein; transcription regulation; zinc finger  
 E:58-94/Region: zinc finger CCCC motif

Query Match	80.4%	Score 74	DB 1	Length 98
Best Local Similarity	76.5%	Pred. No.	0.00035	
Matches 13	Conservative	2	Mismatches	2
			Indels	0
			Gaps	0
Qy	1	LLMGLTIVCPICSOKE	17	
Db	82	LLGLTINIVCPICAPKP	98	

RESULT 3  
M7WL35  
E7 protein - human papillomavirus type 35  
C:Species: human papillomavirus type 35  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Mar-2001  
C:Accession: F40824; S36522  
R:Marich, J.E.; Pontier, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.  
Virology 186, 770-776, 1992  
A:Title: The phylogenetic relationship and complete nucleotide sequence of human papillomavirus type 35  
A:Reference number: F40824; MUID:92124753  
A:Accession: F40824  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-99 <MAR>  
A:Cross-references: GB:M74117; NID:g333050; PIDN:AAA46967.1; PID:g333052  
R:Deilus, H.; Hofmann, B.  
submitted to the EMBL Data Library, August 1993  
A:Description: Primer-directed sequencing of human papillomavirus types.  
A:Reference number: S36469  
A:Accession: S36522  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-99 <DEL>  
A:Cross-references: EMBL:X74477; NID:g396997; PIDN:CAAS2562.1; PID:g396999  
A:Experimental source: strain 35H  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation; zinc finger  
A:59-95/Region: zinc finger CCCC motif

```

Query Match          76.1%; Score 70; DB 1; Length 99;
Best Local Similarity 81.2%; Pred. No. 0.0014;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1 LLMGTLGIVCPICCSOK 16
          ||||| ||||| |||
Db       83 LLMGTFGIVCPGCCSOR 98

RESULT  4
W7WL6
E7 protein - human papillomavirus type 6b
C:Species: human papillomavirus type 6b
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: D20558
R:Schwarz, E.; Durst, M.; Demanowski, C.; Lattemann, O.; Zech, R.; Wolfspenger, E.; Sch
EMBO J. 2, 2341-2348, 1983
A:Title: DNA sequence and genome organization of genital human papillomavirus type 6b.
:Reference number: A90975; MUID:84131949

```

A:Accession: J20558  
A:Molecule\_type: DNA  
A:Residues: 1..98 <SCH>  
A:Cross-references: CB:K00203; NID:g60955; PID:CAA25019.1; PID:g60957  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation; zinc finger  
F:58-94/Region: zinc finger CCCC motif

Query Match	72.8%;	Score 67;	DB 1;	Length 98;
Best Local Similarity	75.0%;	Pred. No. 0.0039;		
Matches 12;	Conservative 2;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      1 LLMGTGIVCPICSQK 16
      ||:|||||
Db      82 LLLGTLNIVCPICAPK 97
```

RESULT 5 E7 protein - human papillomavirus type 13  
W7ML13  
C:Species: human papillomavirus type 13  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1993 #sequence.revision 30-Jun-1993 #text.change 16-Jul-1999  
C:Accession: B42955  
R:van Ransst, M.; Fuse, A.; Flten, P.; Beukem, E.; Pfister, H.; Burk, R.D.; Opdenakker  
Virology 190, 587-596, 1992  
A:Title: Human papillomavirus type 13 and p7gmy chimpanzee papillomavirus type 1: Com  
A:Reference number: A42955; MUID:92391075  
A:Accession: B42955  
A:Molecule type: DNA  
A:Residues: 1-101 <VAN>  
A:Cross-references: EMBL:X62843; MID:960295; PIDN:CAA44648.1; PID:960297  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein  
C:61-97/Region: zinc finger CCCC motif  
C:61-97/Region: zinc finger CCCC motif

Query Match 70.7%; Score 65; DB 1; Length 101;  
Best Local Similarity 68.8%; Pred. No. 0.0079;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
QY      1 LLMGTLGIVCPICSR 16
         ||:|||||:|
Db      85 LLGTLNIVCPICAPK 100
```

```

RESULT      6
S19907
E7-C protein - human papillomavirus type 33
C:Species: human papillomavirus type 33
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 26-Aug-1999
C:Accession: S19907
R:Smithders, P.-J.F.; van den Brule, A.-J.C.; Schrijnemakers, H.F.J.; Raaphorst, P.M.C.;
submitted to the EMBL Data Library, January 1992
A:Description: HPV type 33 in a tonsillar carcinoma generates its putative E7 mRNA v1
A:Reference number: S19906
A:Accession: S19907
A:Molecule type: mRNA
A:Residues: 1-55 <SNT>
A:Cross-references: EMBL:X64086; NID:g60282; PIDD:CAA5436.1; PID:g60284
C:Superfamily: papillomavirus E7 protein
:Keywords: early protein

```

Query Match	69.68;	Score 64;	DB 2;	Length 55;
Best Local Similarity	68.88;	Pred. No. 0.0067;		
Matches 11;	Conservative 3;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      1 LMGTLGIVCPICSQK 16
         ||||:||||1:1:
Db      40 LMGTVNIVCPTCAQK 55
```

RESULT 7  
W7ML33  
E7 protein - human papillomavirus type 33  
C:Species: human papillomavirus type 33  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999  
C:Accession: A03689; S23831; S23827  
R:Cole, S.T.; Strecek, R.E.  
J. Virol. 58, 991-995, 1986  
A:Title: Genome organization and nucleotide sequence of human papillomavirus type 33, with  
A:Reference number: A93020; MUID:86200464  
A:Accession: A03689  
A:Molecule type: DNA  
A:Residues: 1-97 <COL>  
A:Cross-references: GB:M12737; NID:9333049; PIDN:AAA46959.1; PID:9463178  
R:Snijders, P.J.F.; van den Brule, A.J.C.; Schrijnemakers, H.F.J.; Raaphorst, P.M.C.; Me  
submitted to the EMBL Data Library, January 1992  
A:Description: HPV type 33 in a tonsillar carcinoma generates its putative E7 mRNA via t  
A:Reference number: S19906  
A:Accession: S23831  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-97 <SNT>  
A:Cross-references: EMBL:X64085; NID:960278; PIDN:CAA45434.1; PID:960281; EMBL:X64084; N  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation; zinc finger  
F:58-94/Region: zinc finger CCCC motif

Query Match 69.6%; Score 64; DB 1; Length 97;  
Best Local Similarity 68.8%; Pred. No. 0.011;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLMGTGIVCPICSOX 16  
|||||:|||||:|:  
Db 82 LLMGTIVIVCPICAO 97

RESULT 8  
W7ML31  
E7 protein - human papillomavirus type 31  
C:Species: human papillomavirus type 31  
A:Note: host Homo sapiens (man)  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 16-Jul-1999  
C:Accession: B32444  
R:Goldsbrough, M.D.; Disilvestre, D.; Temple, G.F.; Lorincz, A.T.  
Virology 171, 306-311, 1989  
A:Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-asso  
A:Reference number: A94398; MUID:89299478  
A:Accession: B32444  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-98 <COL>  
A:Cross-references: GB:J04353; NID:9333048; PIDN:AAA46951.1; PID:9459917  
C:Comment: This protein may be involved in the oncogenic potential of this virus.  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation; zinc finger  
F:58-94/Region: zinc finger CCCC motif

Query Match 66.3%; Score 61; DB 1; Length 98;  
Best Local Similarity 68.8%; Pred. No. 0.03;  
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LLMGTGIVCPICSOX 16  
|||||:|||||:|:  
Db 82 LLMGTGIVCPICSTR 97

RESULT 9  
W7MLR1  
E7 protein - rhesus papillomavirus (type 1)  
C:Species: rhesus papillomavirus

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Feb-1997  
C:Accession: B38503  
R:Ostrow, R.S.; Labresh, K.V.; Faras, A.J.  
Virology 181, 424-429, 1991  
A:Title: Characterization of the complete RHPV 1 genomic sequence and an integration  
A:Reference number: A38503; MUID:91135018  
A:Accession: B38503  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-113 <OST>  
A:Cross-references: EMBL:M37717  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 66.3%; Score 61; DB 1; Length 113;  
Best Local Similarity 68.8%; Pred. No. 0.034;  
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LLMGTGIVCPICSOX 16  
|||||:|||||:|:  
Db 97 LLMGTIDIVCPICASR 112

RESULT 10  
W7ML42  
E7 protein - human papillomavirus type 42  
C:Species: human papillomavirus type 42  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Feb-1997  
C:Accession: F39451  
R:Philipp, W.; Honore, N.; Sapp, M.; Cole, S.T.; Strecek, R.E.  
Virology 186, 331-334, 1992  
A:Title: Human papillomavirus type 42: new sequence, conserved genome organization.  
A:Reference number: A39451; MUID:92087479  
A:Accession: F39451  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-93 <PHI>  
A:Cross-references: GB:M73236  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 64.1%; Score 59; DB 1; Length 93;  
Best Local Similarity 60.0%; Pred. No. 0.058;  
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLMGTGIVCPICSO 15  
|||||:|||||:|:  
Db 77 MLGTIDIVCPICAR 91

RESULT 11  
W7ML58  
E7 protein - human papillomavirus type 58  
C:Species: human papillomavirus type 58  
A:Note: host Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 21-Jul-2000  
C:Accession: F36779  
R:Kiril, Y.; Iwamoto, S.; Matsukura, T.  
Virology 185, 424-427, 1991  
A:Title: Human papillomavirus type 58 DNA sequence.  
A:Reference number: A36779; MUID:92024102  
A:Accession: F36779  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-98 <KIR>  
A:Cross-references: GB:D90400; NID:9222386; PIDN:BAA31846.1; PID:93337099  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation; zinc finger  
F:59-95/Region: zinc finger CCCC motif

Query Match 64.1%; Score 59; DB 1; Length 98;  
 Best Local Similarity 68.8%; Pred. No. 0.06;  
 Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLMGTLGIVCPICSK 16  
 ||||| ||||| :  
 Db 83 LLMGCTCTVCPSCAQ 98

## RESULT 12

S36585  
 E7 protein - human papillomavirus type 7

C:Species: human papillomavirus type 7

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36585

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36585

A:Molecule type: DNA

A:Residues: 1-111 <DEL>

A:Cross-references: EMBL:X74463; NID:9397060; PIDN:CAA52477.1; PID:9397062

C:Superfamily: papillomavirus E7 protein

C:Keywords: DNA binding; early protein; transcription regulation

Query Match 64.1%; Score 59; DB 2; Length 111;  
 Best Local Similarity 78.6%; Pred. No. 0.067;  
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMGTLGIVCPICS 14  
 ||||| ||||| :  
 Db 95 LLMGTLNIVCPNCA 108

## RESULT 13

S36516  
 E7 protein - human papillomavirus type 34

C:Species: human papillomavirus type 34

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36516

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36516

A:Molecule type: DNA

A:Residues: 1-97 <DEL>

A:Cross-references: EMBL:X74476; NID:9396989; PIDN:CAA52556.1; PID:9396991

C:Superfamily: papillomavirus E7 protein

C:Keywords: DNA binding; early protein; transcription regulation

Query Match 63.0%; Score 58; DB 2; Length 97;  
 Best Local Similarity 68.8%; Pred. No. 0.084;  
 Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLMGTLGIVCPICSK 16  
 ||||| ||||| :  
 Db 81 LLMGALKIVCPNCR 96

## RESULT 14

S36510  
 E7 protein - human papillomavirus type 32

C:Species: human papillomavirus type 32

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36510

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469  
 A:Accession: S36510  
 A:Molecule type: DNA  
 A:Residues: 1-104 <DEL>  
 A:Cross-references: EMBL:X74475; NID:9396981; PIDN:CAA52550.1; PID:9396983  
 C:Superfamily: papillomavirus E7 protein  
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 63.0%; Score 58; DB 2; Length 104;  
 Best Local Similarity 64.3%; Pred. No. 0.089;  
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLMGTLGIVCPICS 14  
 ||: ||||| :  
 Db 88 MLMPLTGLIVCPICA 101

## RESULT 15

W7WLC1

E7 protein - pygmy chimpanzee papillomavirus (type 1)

C:Species: pygmy chimpanzee papillomavirus

C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Jul-1999

C:Accession: B36818

R:van Ranst, M.; Fuse, A.; Eiten, P.; Beuken, E.; Pfister, H.; Burk, R.D.; Opdenakker

Virology 190, 587-596, 1992

A:Title: Human papillomavirus type 13 and pygmy chimpanzee papillomavirus type 1: Com

A:Reference number: A42955; MUID:92391075

A:Accession: B36818

A:Molecule type: DNA

A:Residues: 1-98 <VAN>

A:Cross-references: EMBL:X62844; NID:961010; PIDN:CAA44656.1; PID:961012

C:Superfamily: papillomavirus E7 protein

C:Keywords: DNA binding; early protein; transcription regulation; transforming protel

F:58-94/Region: zinc finger CCCC motif

Query Match 62.0%; Score 57; DB 1; Length 98;  
 Best Local Similarity 64.3%; Pred. No. 0.12;  
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLMGTLGIVCPICS 14  
 ||: ||||| :  
 Db 82 LLMGSLNIVCPICA 95

Search completed: February 13, 2002, 10:00:43  
 Job time: 377 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 10:02:20 : Search time 21.6 seconds

(without alignments)  
28.857 Million cell updates/sec

Title: US-09-664-225-69

Perfect score: 92

Sequence: 1 LMGTGIVCPICSKP 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	98	VE7_HPV16	P03129 human papill
2	74	80.4	98	VE7_HPV11	P04020 human papill
3	70	76.1	99	VE7_HPV35	P27330 human papill
4	67	72.8	98	VE7_HPV6B	P06464 human papill
5	66	71.7	98	VE7_HPV6A	O84292 human papill
6	65	70.7	101	VE7_HPV13	O02272 human papill
7	64	69.6	97	VE7_HPV33	P06429 human papill
8	61	66.3	98	VE7_HPV31	P17387 human papill
9	61	66.3	113	VE7_HPV1	P22161 rhesus papill
10	59	64.1	93	VE7_HPV42	P27331 human papill
11	59	64.1	98	VE7_HPV58	P26557 human papill
12	58	64.1	111	VE7_HPV07	P36816 human papill
13	58	63.0	97	VE7_HPV34	P36828 human papill
14	58	63.0	97	VE7_HPV44	O80914 human papill
15	58	63.0	104	VE7_HPV32	P36827 human papill
16	57	62.0	97	VE7_HPV55	O80935 human papill
17	57	62.0	98	VE7_HPV40	O02272 pygmy chimp
18	57	62.0	111	VE7_HPV40	P36829 human papill
19	56	60.9	105	VE7_HPV53	P36832 human papill
20	55	59.8	88	VE7_HPV63	O07958 human papill
21	54	58.7	105	VE7_HPV30	P36826 human papill
22	52	56.5	87	VE7_HPV28	P50783 human papill
23	52	56.5	105	VE7_HPV56	P36833 human papill
24	52	56.5	105	VE7_HPV66	O80956 human papill
25	51	55.4	86	VE7_HPV10	P36818 human papill
26	51	55.4	99	VE7_HPV52	P36831 human papill
27	51	55.4	101	VE7_HPV51	P26558 human papill
28	49	53.3	93	VE7_HPV1A	P06465 human papill
29	49	53.3	95	VE7_HPV61	O80949 human papill
30	48	52.2	95	VE7_HPV17	P36821 human papill
31	46	50.0	90	VE7_HPV29	P50784 human papill
32	46	50.0	109	VE7_HPV39	P24637 human papill
33	45	48.9	708	FSYA_DROME	P25028 drosophila

## ALIGNMENTS

RESULT	1	STANDARD:	PRT:	98 AA.
VE7_HPV16				
ID	VE7_HPV16			
AC	P03129;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	E7 PROTEIN.			
OS	Human papillomavirus type 16.			
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			
OX	Papillomavirus.			
NCBI_TaxID=10581;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=65246220; PubMed=2990099;			
RA	Seedorf K., Kitzmer G., Durst M., Suhai S., Rowekamp W.G.;			
RT	"Human papillomavirus type 16 DNA sequence.";			
RL	Virology 145:181-185(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90218027; PubMed=2157796;			
RA	Schneider-Munoury S., Penau-Annaudet G., Breitburd F., Orth G.;			
RT	"Expression of the human papillomavirus type 16 genome in SK-V cells,			
RL	a line derived from a vulvar intraepithelial neoplasia.";			
RN	J. Gen. Virol. 71:809-817(1990).			
RP	SEQUENCE FROM N.A.			
RA	Song Y.S., Kee S.H., Kim J.W., Park N.H., Kang S.B., Lee H.P.;			
RT	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.			
RL	[4]			
RP	SEQUENCE FROM N.A.			
RA	Tornesello M.L., Buongiorno F.M., Meglio A., Buongiorno L.,			
RT	Beth-Giraldo E., Giraldo G.; to the EMBL/GenBank/DBJ databases.			
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	FUNCTION.			
RX	MEDLINE=88223347; PubMed=2836062;			
RA	Phelps W.C., Yee C.L., Munger K., Howley P.M.;			
RT	"The human papillomavirus type 16 E7 gene encodes transactivation and			
RL	transformation functions similar to those of adenovirus E1A.";			
Cell 53:539-547(1988).				
CC	-1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING			
CC	ACTIVITIES.			
CC	-1- DISEASE: HPV16, IN COMPARISON TO HPV TYPES 6 AND 11, IS MORE			
CC	OFTEN ASSOCIATED WITH MALIGNANT GENITAL CANCERS IN HUMANS.			
CC	-1- SIMILARITY: LOCAL WITH ADENOVIRUS E1A AND SV40 LT.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			

34	45	48.9	1496	1	CA25_HUMAN	P05997 homo sapien
35	44	47.8	94	1	VE7_HPV37	O80901 human papill
36	44	47.8	125	1	YBAN_ECOLI	P45808 escherichia
37	44	47.8	252	1	YCIK_ECOLI	P31808 escherichia
38	44	47.8	365	1	MLTA_ECOLI	P46885 escherichia
39	43	46.7	92	1	VE7_HPV57	P22160 human papill
40	43	46.7	93	1	VE7_HPV09	P36817 human papill
41	43	46.7	105	1	VE7_HPV18	P06788 human papill
42	43	46.7	109	1	VE7_HPV70	P50785 human papill
43	43	46.7	149	1	NPc2_BRARE	O90913 brachydanio
44	43	46.7	449	1	Y753_SYNY3	P74635 synchocyst
45	42.5	46.2	261	1	RNG6_MOUSE	P28078 mus muscult

DR EMBL: K02718; AAAA6940.1; -;  
DR EMBL: D00735; BAA00633.1; -;  
DR EMBL: U76411; AAB18962.1; -;  
DR EMBL: U76412; AAB18963.1; -;  
DR EMBL: U76413; AAB18964.1; -;  
DR EMBL: AF003020; AAB70737.1; -;  
DR EMBL: AF003023; AAB70740.1; -;  
DR EMBL: AF003024; AAB70741.1; -;  
DR EMBL: AF003025; AAB70742.1; -;  
DR EMBL: AF003026; AAB70743.1; -;  
DR PIR: A03688; W7MLH5;  
DR InterPro: IPR000148; PapvL\_E7.  
DR Pfam: PF00527; E7; 1.  
KW Early protein; Transcription regulation; Transforming protein;  
FT DOMAIN 58 61 C-XX-C MOTIF 1.  
SQ SEQUENCE 98 AA: 11022 MW: 98D612534CD2C9EB CRC64;

Query Match 100.0%; Score 92; DB 1; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2, 1e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTLGIVPCISQKP 17  
Db 82 LLMGTLGIVPCISQKP 98  
|||||

RESULT 2  
VE7\_HPV11 STANDARD; PRT; 98 AA.  
ID VE7\_HPV11  
AC P04020;  
DT 23-OCT-1986 (Rel. 02, Created)  
DT 23-OCT-1986 (Rel. 02, Last sequence update)  
DT 31-OCT-1996 (Rel. 34, Last annotation update)  
DE E7 PROTEIN.  
GN E7.  
OS Human papillomavirus type 11.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_Taxid=10580;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86181601; Pubmed=3089427;  
RA Dartmann K., Schwarz E., Gissmann L., Zur Hausen H.;  
RT "The nucleotide sequence and genome organization of human papilloma  
virus type 11.";  
RL Virology 151:124-130(1986).  
CC [2]  
CC SEQUENCE FROM N.A.  
RA File K.H., Fan L., Fritsch M.H., Bryan J., Brown D.R.;  
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
CC ACTIVITIES.  
CC -1- SIMILARITY: LOCAL WITH ADENOVIRUS E1A AND SV40 LT.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: M4119; AAA46928.1; -;  
DR EMBL: L36108; AAA21704.1; -;  
DR PIR: A03690; W7MLI1.  
DR InterPro: IPR000148; PapvL\_E7.  
DR Pfam: PF00527; E7; 1.  
KW Early protein; Transcription regulation; Transforming protein;  
FT DOMAIN 58 61 C-XX-C MOTIF 1.  
SQ SEQUENCE 98 AA: 11022 MW: 98D612534CD2C9EB CRC64;

FT DOMAIN 91 94 C-XX-C MOTIF 2.  
SQ SEQUENCE 98 AA: 10889 MW: AACCA9A60C933E1F6 CRC64;

Query Match 80.4%; Score 74; DB 1; Length 98;  
Best Local Similarity 76.5%; Pred. No. 0.00011;  
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLMGTLGIVPCISQKP 17  
Db 82 LLMGTLGIVPCISQKP 98  
|||||

RESULT 3  
VE7\_HPV35 STANDARD; PRT; 99 AA.  
ID VE7\_HPV35  
AC P27230;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE E7 PROTEIN.  
GN E7.  
OS Human papillomavirus type 35.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_Taxid=10587;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94265501; Pubmed=8205838;  
RA Deltus H., Hofmann B.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
CC [2]  
CC SEQUENCE FROM N.A.  
RX MEDLINE=92124753; Pubmed=1310198;  
RA Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;  
RT "The phylogenetic relationship and complete nucleotide sequence of  
human papillomavirus type 35.";  
RL Virology 186:770-776(1992).  
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
CC ACTIVITIES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: M74117; AAA46967.1; -;  
DR EMBL: X74477; CAA52562.1; -;  
DR PIR: F40824; W7MLJ5.  
DR PIR: S36522; S36522.  
DR InterPro: IPR000148; PapvL\_E7.  
DR Pfam: PF00527; E7; 1.  
KW Early protein; Transcription regulation; Transforming protein;  
FT DOMAIN 59 62 C-XX-C MOTIF 1.  
SQ SEQUENCE 99 AA: 10954 MW: BEAABR248051E226 CRC64;

Query Match 76.1%; Score 70; DB 1; Length 99;  
Best Local Similarity 81.2%; Pred. No. 0.00046;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLMGTLGIVPCISQKP 16  
Db 83 LLMGTLGIVPCISQKP 98  
|||||

RESULT 4

```

VE7_HPV6B
ID VE7_HPV6B STANDARD; PRT; 98 AA.
AC P0464;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 6b.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
NCBI_TaxID=10600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84131949; PubMed=6321162;
RA Schwarz E., Durst M., Demankowski C., Laternann O., Zech R.,
RA Wolfberger E., Sinal S., Zur Hausen H.;
RT "DNA sequence and genome organization of genital human papillomavirus
RT type 6b."
RL EMBL J. 2:2341-2348(1983).
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X00203; CAA25019.1; -.
CC DR PIR: D20558; W7ML6.
CC DR InterPro: IPR000148; Papv1_E7.
CC PFam: PF00527; E7; 1.
CC "Early protein: Transcription regulation: Transforming protein;
CC DNA-binding; Trans-acting factor."
CC KW DOMAIN 58 61 C-XX-C MOTIF 1.
CC FT DOMAIN 91 94 C-XX-C MOTIF 2.
CC SQ SEQUENCE 98 AA; 10887 MW; 997CB5095A39C4CD CRC64;

Query Match 72.8%; Score 67; DB 1; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.0013;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMGTGIVCPICSOX 16
Db 82 LLGTLDIVCPICAPK 97

RESULT 5
VE7_HPV6A STANDARD; PRT; 98 AA.
ID VE7_HPV6A
AC Q84292;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 6a.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
NCBI_TaxID=37122;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95297152; PubMed=7778283;
RA Hofmann K.J., Cook J.C., Joyce J.G., Brown D.R., Schultz L.D.,
RA Hofmann K.J., Cook J.C., Joyce J.G., Brown D.R., Schultz L.D.,
RA Hofmann K.J., Cook J.C., Joyce J.G., Brown D.R., Schultz L.D.,
RA "Sequence determination of human papillomavirus type 6a and assembly
RT of virus-like particles in Saccharomyces cerevisiae."
RT Virology 209:506-518(1995).
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.

```

```

CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L41216; AAA74212.1; -.
CC DR InterPro: IPR000148; Papv1_E7.
CC DR PFam: PF00527; E7; 1.
CC KW Early protein: Transcription regulation; Transforming protein;
CC DNA-binding; Trans-acting factor.
CC FT DOMAIN 58 61 C-XX-C MOTIF 1.
CC FT DOMAIN 91 94 C-XX-C MOTIF 2.
CC SQ SEQUENCE 98 AA; 10888 MW; 997CBDB5A39C4CD CRC64;

Query Match 71.7%; Score 66; DB 1; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.0018;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMGTGIVCPICSOX 16
Db 82 LLGTLDIVCPICAPK 97

RESULT 6
VE7_HPV13 STANDARD; PRT; 101 AA.
ID VE7_HPV13
AC Q02271;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 13.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
NCBI_TaxID=10573;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92391075; PubMed=1325697;
RA van Ranst M., Fuse A., Flten P., Beuken E., Flister H., Burk R.D.,
RA Opdenakker G.;
RT "Human papillomavirus type 13 and pygmy chimpanzee papillomavirus
RT type 1: comparison of the genome organizations."
RT Virology 190:587-596(1992).
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X62843; CAA4648.1; -.
CC DR PIR: B42955; W7ML13.
CC DR InterPro: IPR000148; Papv1_E7.
CC DR PFam: PF00527; E7; 1.
CC KW Early protein: Transcription regulation; Transforming protein;
CC DNA-binding; Trans-acting factor.
CC FT DOMAIN 61 64 C-XX-C MOTIF 1.
CC FT DOMAIN 94 97 C-XX-C MOTIF 2.
CC SQ SEQUENCE 101 AA; 11048 MW; 380E5D4B2B0472 CRC64;

Query Match 70.7%; Score 65; DB 1; Length 101;

```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M60184; AAA79312.1; -  
DR PIR; B38503; W7MLR1.  
DR InterPro: IPR000148; Papyl\_E7.  
DR Pfam; PF00527; E7; 1.  
KW Early protein; Transcription regulation; Transforming protein;  
KW DNA-binding; Trans-acting factor.  
FT DOMAIN 73 76 C-XX-C MOTIF 1.  
FT DOMAIN 106 109 C-XX-C MOTIF 2.  
SQ SEQUENCE 113 AA; 12818 MW; 9C404BDAC1298306 CRC64;

Query Match 66.3%; Score 61; DB 1; Length 113;  
Best Local Similarity 68.8%; Pred. No. 0.012;

Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LMGTGLGIVCPICSQK 16  
DB 97 LMGTGLGIVCPICSCAR 112

RESULT 10  
VE7\_HP42 STANDARD; PRT; 93 AA.  
ID VE7\_HP42  
AC P27231;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE E7 PROTEIN.

OS Human papillomavirus type 42.  
OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10590;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92087479; PubMed=1309278;  
RA Philipp W., Honore N., Sapp M., Cole S.T., Streck R.E.;  
RT "Human papillomavirus type 42: new sequences, conserved genome  
RT organization."  
RL Virology 186:331-334(1992).  
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
CC ACTIVITIES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; W73236; AAA7042.1; ALT\_INIT.  
DR PIR; F39451; W7MLA2.  
DR InterPro: IPR000148; Papyl\_E7.  
DR Pfam; PF00527; E7; 1.  
KW Early protein; Transcription regulation; Transforming protein;  
KW DNA-binding; Trans-acting factor.  
FT DOMAIN 53 56 C-XX-C MOTIF 1.  
FT DOMAIN 86 89 C-XX-C MOTIF 2.  
SQ SEQUENCE 93 AA; 10679 MW; 829DIE24071F7AE3 CRC64;

Query Match 64.1%; Score 59; DB 1; Length 93;  
Best Local Similarity 60.0%; Pred. No. 0.02;  
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 LMGTGLGIVCPICSQ 15  
DB 77 LMGTGLGIVCPICAR 91

RESULT 11  
VE7\_HP58 STANDARD; PRT; 98 AA.  
ID VE7\_HP58  
AC P26557;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE E7 PROTEIN.

GN E7.  
OS Human papillomavirus type 58.  
OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10598;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92024102; PubMed=1656594.  
RA Kiril V., Iwamoto S., Matsukura T.;  
RT "Human papillomavirus type 58 DNA sequence."  
RL Virology 185:424-427(1991).  
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
CC ACTIVITIES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D90400; BAA31846.1; -  
DR PIR; F36779; W7ML58.  
DR InterPro: IPR000148; Papyl\_E7.  
DR Pfam; PF00527; E7; 1.  
KW Early protein; Transcription regulation; Transforming protein;  
KW DNA-binding; Trans-acting factor.  
FT DOMAIN 59 62 C-XX-C MOTIF 1.  
FT DOMAIN 92 95 C-XX-C MOTIF 2.  
SQ SEQUENCE 98 AA; 10819 MW; D719984F3CF48D12 CRC64;

Query Match 64.1%; Score 59; DB 1; Length 98;  
Best Local Similarity 68.8%; Pred. No. 0.021;  
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LMGTGLGIVCPICSQ 16  
DB 83 LMGTGLGIVCPICCAO 98

RESULT 12  
VE7\_HP07 STANDARD; PRT; 111 AA.  
ID VE7\_HP07  
AC P36816;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE E7 PROTEIN.

GN E7.  
OS Human papillomavirus type 7.  
OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10620;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94265501; PubMed=8205838;  
RA Delius H., Hofmann B.;

RP "Primer-directed sequencing of human papillomavirus types."  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
 CC ACTIVITIES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X74463; CAAS2477.1; -.  
 DR PIR: S36585; S36585.  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 KW Early protein; Transcription regulation; Transforming protein;  
 KW DNA-binding; Trans-acting factor.  
 FT DOMAIN 71 74 C-XX-C MOTIF 1.  
 FT DOMAIN 104 107 C-XX-C MOTIF 2.  
 SO SEQUENCE 111 AA; 12460 MW; 830ADCA5F912BF1A CRC64;

Query Match 64.1%; Score 59; DB 1; Length 111;  
 Best Local Similarity 78.6%; Pred. No. 0.023;  
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLMGLTIVCPICS 14  
 Db 95 LLMGLTIVCPICA 108

RESULT 13  
 ID VE7\_HPV34 STANDARD: PRT: 97 AA.  
 AC P36828;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 34.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10613;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94265501; PubMed=8205838;  
 RA Dellus H.; Hofmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types."  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
 CC ACTIVITIES.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: X74476; CAAS2556.1; -.  
 DR PIR: S36516; S36516.  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 KW Early protein; Transcription regulation; Transforming protein;  
 KW DNA-binding; Trans-acting factor.  
 FT DOMAIN 57 60 C-XX-C MOTIF 1.  
 FT DOMAIN 90 93 C-XX-C MOTIF 2.  
 SO SEQUENCE 97 AA; 10985 MW; 19E3C9D1F12BF4F6 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 97;  
 Best Local Similarity 68.8%; Pred. No. 0.029;  
 Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LLMGLTIVCPICSOK 16  
 Db 81 LLMGLTIVCPICARR 96

RESULT 14  
 ID VE7\_HPV44 STANDARD: PRT: 97 AA.  
 AC Q80914;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 44.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10592;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dellus H.;  
 RT Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
 CC ACTIVITIES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U31788; AAA79458.1; -.  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 KW Early protein; Transcription regulation; Transforming protein;  
 KW DNA-binding; Trans-acting factor.  
 FT DOMAIN 57 60 C-XX-C MOTIF 1.  
 FT DOMAIN 90 93 C-XX-C MOTIF 2.  
 SO SEQUENCE 97 AA; 10641 MW; E4866AE13B050456 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 97;  
 Best Local Similarity 56.2%; Pred. No. 0.029;  
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLMGLTIVCPICSOK 16  
 Db 81 LLMGLTIVCPICARR 96

RESULT 15  
 ID VE7\_HPV32 STANDARD: PRT: 104 AA.  
 AC P36827;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 32.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10612;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94265501; PubMed=8205838;

```

RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74475; CAA52550.1; -
DR PIR: S36510; S36510.
DR InterPro: IPR000148; Papyl_E7.
DR Pfam: PF00527; E7; 1.
KW Early protein; Transcription regulation; Transforming protein;
KW DNA-binding; Trans-acting factor.
FT DOMAIN 64 67 C-XX-C MOTIF 1.
FT DOMAIN 97 100 C-XX-C MOTIF 2.
SQ SEQUENCE 104 AA; 11591 MW; F9DFRC086D804FF CRC64;

```

```

Query Match          63.0%; Score 58; DB 1; Length 104;
Best Local Similarity 64.3%; Pred. No. 0.031;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 LLMGTGLGIVCPICS 14
   :1:1111111111:
DB 88 MLDPTGLGIVCPICA 101

```

Search completed: February 13, 2002, 10:02:20  
Job time: 249 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 13, 2002, 10:01:52 ; Search time 62.52 Seconds  
(without alignments)  
39.773 Million cell updates/sec

Title: US-09-664-225-69

Perfect score: 92

Sequence: 1 LMGTLGIVPCISQKP 17

Scoring table:

BLOSUM62

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP invertebrate: \*  
6: SP mammal: \*  
7: SP mhc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP virus: \*  
13: SP vertebrate: \*  
14: SP unclassified: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	26	12	Q84267 human papill
2	92	100.0	98	12	Q12337 human papill
3	92	100.0	98	12	Q12338 human papill
4	92	100.0	98	12	Q11650 human papill
5	69	75.0	97	12	Q82006 human papill
6	67	72.8	98	12	Q90LP4 human papill
7	65	70.7	99	12	Q90724 human papill
8	64	69.6	55	12	Q81886 human papill
9	63	68.5	96	12	Q98005 human papill
10	62	67.4	93	12	Q90DH8 human papill
11	62	67.4	93	12	Q90DH6 human papill
12	62	67.4	93	12	Q9QDH4 human papill
13	62	67.4	93	12	Q9QDH2 human papill
14	59	64.1	98	12	Q90CZ2 human papill
15	59	62.0	98	12	Q90CZ1 human papill
16	57	62.0	99	12	Q90CZ1 human papill
17	52	56.5	336	10	Q9C7E9 common chlam
18	50	54.3	108	12	Q9WHG1 arabidopsis
19	49	53.3	88	5	Q76959 manduca sex

20	49	53.3	107	12	Q81965 human papill
21	48	52.2	96	12	Q84355 mastomys na
22	46	50.0	90	12	Q56947 human papill
23	46	50.0	100	12	Q91R58 human papill
24	46	50.0	228	4	Q91508 human papill
25	46	50.0	228	11	Q9ET26 human papill
26	46	50.0	354	10	Q9CJ26 mus musculu
27	45	48.9	93	12	Q56953 human papill
28	45	48.9	195	5	Q9V921 drosophila
29	45	48.9	1006	2	P73012 synechocyst
30	44	47.8	100	12	Q81998 human papill
31	44	47.8	354	5	Q61673 aplysia cal
32	44	47.8	388	2	Q52176 myxococcus
33	44	47.8	449	2	Q9CGJ0 lactococcus
34	44	47.8	458	2	Q9XB07 enterococcus
35	44	47.8	462	11	Q9CXM8 mus musculu
36	44	47.8	580	2	Q9F861 enterococu
37	44	47.8	580	4	Q9H8W5 homo sapien
38	44	47.8	639	5	Q61858 caenorhabdi
39	44	47.8	723	11	Q35193 mus musculu
40	44	47.8	852	11	Q35192 mus musculu
41	43.5	47.3	93	2	Q9Z8G9 chlamydia p
42	43.5	47.3	525	5	Q76235 trypanosoma
43	43.5	47.3	525	5	Q9NCZ9 trypanosoma
44	43	46.7	129	2	Q9AFU4 shigella fl
45	43	46.7	145	4	Q9H509 homo sapien

#### ALIGNMENTS

RESULT	ID	Q84267	PRELIMINARY:	PRT:	26 AA.
AC	Q84267				
DT	01-NOV-1996	(TREMBLrel. 01, Created)			
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)			
DE	E7 ORF (FRAGMENT).				
OS	Human papillomavirus.				
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;				
CC	Papillomavirus.				
OX	NCBI_Taxid=10566;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=6905007; PubMed=2536104;				
RA	Choo K.B., Cheung W.F., Liew L.N., Lee H.H., Han S.H.;				
RT	"Presence of catenated human papillomavirus type 16 episomes in a				
RT	cervical carcinoma cell line."				
RL	J. Virol. 63:782-789(1989).				
DR	EMBL: M24215; AAA46944.1; -				
DR	InterPro: IPR000148; PapVI_E7.				
DR	Pfam: PF00527; E7; 1.				
FT	NON_TER				
SQ	SEQUENCE	26 AA: 2799 MW: 91C16FD134D18B34 CRC64;			
Query Match					
Best Local Similarity 100.0%; Score 92; DB 12; Length 26;					
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
OY	1	LMGTLGIVPCISQKP 17			
DB	10	LMGTLGIVPCISQKP 26			
RESULT 2					
ID	Q12337	PRELIMINARY:	PRT:	98 AA.	
AC	Q12337				
DT	01-JUL-1997	(TREMBLrel. 04, Created)			
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)			
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)			

DE E7 PROTEIN.  
 OS Human papillomavirus type 16.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10581;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SIHA;  
 RX MEDLINE=97437474; PubMed=9292007;  
 RA Tornesello M.L., Buonaguro F.M., Meglio A., Buonaguro L.,  
 RA Beth-Girardo E., Girardo G.;  
 RT "Sequence variations and viral genomic state of human papillomavirus  
 RT type 16 in penile carcinomas from Ugandan patients.";  
 RL J. Gen. Virol. 78:0-0(0).  
 DR EMBL; AF003021; AAB70738.1; -;  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 SQ SEQUENCE 98 AA; 11056 MW; 19DBB8F14CD2C705 CRC64;

Query Match 100.0%; Score 92; DB 12; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTLGIVCPICSOXP 17  
 |||||  
 DB 82 LLMGTLGIVCPICSOXP 98

RESULT 3  
 ID 012338 PRELIMINARY; PRT; 98 AA.  
 AC 012338;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE E7 PROTEIN.  
 OS Human papillomavirus type 16.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10581;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PC-4;  
 RX MEDLINE=97437474; PubMed=9292007;  
 RA Tornesello M.L., Buonaguro F.M., Meglio A., Buonaguro L.,  
 RA Beth-Girardo E., Girardo G.;  
 RT "Sequence variations and viral genomic state of human papillomavirus  
 RT type 16 in penile carcinomas from Ugandan patients.";  
 RL J. Gen. Virol. 78:0-0(0).  
 DR EMBL; AF003022; AAB70739.1; -;  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 SQ SEQUENCE 98 AA; 10969 MW; 9BD612534CCEA59B CRC64;

Query Match 100.0%; Score 92; DB 12; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTLGIVCPICSOXP 17  
 |||||  
 DB 82 LLMGTLGIVCPICSOXP 98

RESULT 4  
 ID 011650 PRELIMINARY; PRT; 98 AA.  
 AC 011650;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE E7 PROTEIN.

GN E7.  
 OS Human papillomavirus type 16.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10581;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AF078347; AFA7689T;  
 RX MEDLINE=97407827; PubMed=9264576;  
 RA Song Y.-S., Kee S.-H., Kim J.-W., Park N.-H., Kang S.-B., Chang W.-H.,  
 RA Lee H.-P.;  
 RT "Major sequence variants in E7 gene of human papillomavirus type 16  
 RT from cervical cancerous and noncancerous lesions of Korean women.";  
 RL Gynecol. Oncol. 66:275-281(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AF078347; AFA7689T;  
 RX van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,  
 RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;  
 RT "Analysis of Human papillomavirus type 16 E6 variants in relation to  
 RT p53 codon 72 polymorphism genotype in cervical carcinogenesis.";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U76404; AAC58243.1; -;  
 DR EMBL; AJ388063; CAB45119.1; -;  
 DR EMBL; AJ388062; CAB45117.1; -;  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 SQ SEQUENCE 98 AA; 10995 MW; 81E53B534CC3281B CRC64;

Query Match 100.0%; Score 92; DB 12; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTLGIVCPICSOXP 17  
 |||||  
 DB 82 LLMGTLGIVCPICSOXP 98

RESULT 5  
 ID 082006 PRELIMINARY; PRT; 97 AA.  
 AC 082006;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE E6, E7, E1, E2, E4, L2, AND L1 GENES.  
 GN E7.  
 OS Human papillomavirus type 7.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10620;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96213783; PubMed=8635859;  
 RA Voelster C., He Y., De Villiers E.M.,  
 RA Greenspan D., De Villiers E.M.;  
 RT "Novel HPV types present in oral papillomatous lesions from patients  
 RT with HIV infection.";  
 RL Int. J. Cancer 66:453-456(1996).  
 DR EMBL; X94165; CAA63883.1; -;  
 DR InterPro: IPR000148; Papv1\_E7.  
 DR Pfam: PF00527; E7; 1.  
 SQ SEQUENCE 97 AA; 10970 MW; 651D0345D048F022 CRC64;

Query Match 75.0%; Score 69; DB 12; Length 97;  
 Best Local Similarity 86.7%; Pred. No. 0.0032;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLMGTLGIVCPICSO 15  
 |||||  
 DB 81 LLMGTLGIVCPICSR 95

RESULT	6			
09QLP4				
ID	09QLP4	PRELIMINARY;	PRT;	98 AA.
AC	09QLP4;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
DE	COMPLETE GENOME.			
GN	E7.			
OS	Human papillomavirus type 6.			
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			
OC	Papillomavirus.			
OX	NCBI_TaxID=31552;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HUMAN PAPILLOMAVIRUS TYPE 6VC;			
RA	Kovelman R., Bilter G.K., Roman A., Brown D.R., Barbosa M.S.;			
RT	"Human papillomavirus type 6: classification of clinical isolates and			
RT	functional analysis of E2 proteins.";			
RL	J. Gen. Virol. 0:0-0(1999).			
LN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HUMAN PAPILLOMAVIRUS TYPE 6VC;			
RA	Kovelman R., Bilter G.K., Roman A., Brown D.R., Barbosa M.S.;			
RL	Submitted (SPP-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF092932; AAF00065.1; "			
DR	InterPro: IPR000148; PapvL_E7.			
DR	Pfam: PF00527; E7; 1.			
SO	SEQUENCE 98 AA; 10903 MW; 74DD65095A39DDE5 CRC64;			

```

Query Match      72.8%   Score 67; DB 12; Length 98;
Best Local Similarity 75.0%   Pred No. 0.0065;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0.
OY      1 LLMGTLGIVCPICSOQ 16
          ||:|||||
db      82 LLLGTLINIVCPICAPK 97

```

```

RESULT 7
090724
ID 090724 PRELIMINARY; PRT; 99 AA.
AC 090724;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE ORF E7.
OS Human papillomavirus type 67.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=37120;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiriti Y., Iwamoto S., Matsukura T.;
RT "Complete DNA sequence of human papillomavirus type 67 isolated from a
RT vaginal intraepithelial neoplasia."
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: D21208; BAA2885.1; -
DR InterPro: IPR000148; Papv1_E7.
DR Pfam: PF00527; E7; 1.
SO SEQUENCE 99 AA; 1156 MW; 81AA636D758BC44 CRC64;

```

Query Match	70.7%	Score 65;	DB 12;	Length 99;
Best Local Similarity	68.8%	Pred. No. 0.013;		
Matches 11;	Conservative 3;	Mismatches 2;	Indels 0;	Gaps 0;
OY	1	ILMGTGIVCPICSO	K	16
		:11		1:1
Db	84	MLMNTGIVCPICAO		99
		MLMNTGIVCPICAO		99

RESULT	8				
ID	Q81886	PRELIMINARY:	PRT:	55 AA.	,
AC	Q81886;				
DT	01-NOV-1996	(TREMBLrel. 01, Created)			
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)			
DE	TYPE 33 POT. B6*II, E7-C, E1^E4.				
CS	E7-C.				
OC	Human papillomavirus type 33.				
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;				
OC	Papillomavirus.				
OX	NCBI_Taxid=10586;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=TYPE 33;				
RC	MEDLINE=92219410; PubMed=1113922;				
RA	Snijders P.J.F., Van den Brule A.J.C., Schrijnemakers H.F.J.,				
RA	Raaphorst P.M.C., Melijer C.J.L.M., Walboomers J.M.N.;				
RT	"Human papillomavirus type 33 in a tonsillar carcinoma generates its				
RT	putative E7 mRNA via two E6* transcript species which are terminated				
RT	at different early region poly(A) sites."				
RL	J. Virol. 66:3172-3178(1992).				
DR	EMBL; X64086; CAA45436.1; -;				
DR	InterPro: IPR000148; Papv1_E7.				
DR	Pfam: PF00527; E7; 1				
SO	SEQUENCE	55 AA; 6246 MW; 8024D6ECCA012247 CRC64;			

Query Match	69.6%	Score 64	DB 12	Length 55
Best Local Similarity	69.8%	Pred. No. 0.011		
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0.				
QY 1 LLMGTGIVCPICSOX 16				
:      :    :				
Db 40 LLMGTVINIVCPICAOO 55				

RESULT 9		PRELIMINARY:		PRT:		96 AA.	
Q98005	Q98005						
AC	Q98005						
DT	01-FEB-1997	(TREMBlrel, 02,	Created)				
DT	01-FEB-1997	(TREMBlrel, 02,	Last sequence update)				
DI	01-JUN-2001	(TREMBlrel, 17,	Last annotation update)				
DE	E7 PROTEIN.						
GN	E7.						
OS	Human papillomavirus type 7.						
OC	Viruses: dsDNA viruses, no RNA stage: Papillomaviridae:						
OC	Papillomavirus.						
OX	NCBI_TaxID=10620:						
RN	111						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=96379050. Pubmed=8784613.						
RA	Longuet M., Cassonnet P., Orth G.:						
RT	A novel genital human papillomavirus (HPV), HPV type 74, found in						
RT	immunosuppressed patients.;						
RL	J. Clin. Microbiol. 34:1859-1862(1996).						
EMBL:	U40822: AAC55128.1: -;						
DR	InterPro: IPR0001048; PapvL_E7.						
DR	Pfam: PF00527; E7; 1.						
SO	SEQUENCE 96 AA; 10522 MW; 80PF8480CEE5CCEN3 CRC64;						

Query March	68.5%	Score 63	DB 12	Length 96
Best Local Similarity	62.5%	Pred. No.	0.025	
Matches 10	Conservative	4	Mismatches	2
			Indels	0
			Gaps	0
Oy	1	LIMGTGIVCPICSO	K	16
		:    :    :		
db	80	LILGTILNIVCPICAP	K	95

RESULT 10  
Q9QDH8 PRELIMINARY; PRT; 93 AA.  
AC Q9QDH8: 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE E7 PROTEIN (FRAGMENT).  
GN E7.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
CC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KE6-1;  
RA Lee H.P., Song Y.S., Kim J.W., Roh J.W., Park N.H., Kang S.B.;  
RT "The most commonly found Human Papillomavirus type 16 E6 and E7  
sequence variation tend to occur together."  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF187866; AAF13393.1; -  
DR InterPro: IPR000148; Papv1\_E7.  
DR Pfam: PF00527; E7; 1.  
FT NON\_TER 93  
SQ SEQUENCE 93 AA: 10452 MW: 83281BB2AE2C9A1F CRC64;

Query Match 67.4%; Score 62; DB 12; Length 93;  
Best Local Similarity 100.0%; Pred. No. 0.034;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTLGIVCPI 12  
|||||  
Db 82 LLMGTLGIVCPI 93

RESULT 11  
Q9QDH6 PRELIMINARY; PRT; 93 AA.  
AC Q9QDH6: 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE E7 PROTEIN (FRAGMENT).  
GN E7.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
CC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KE6-2;  
RA Lee H.P., Song Y.S., Kim J.W., Roh J.W., Park N.H., Kang S.B.;  
RT "The most commonly found Human Papillomavirus type 16 E6 and E7  
sequence variation tend to occur together."  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF187867; AAF13395.1; -  
DR InterPro: IPR000148; Papv1\_E7.  
DR Pfam: PF00527; E7; 1.  
FT NON\_TER 93  
SQ SEQUENCE 93 AA: 10513 MW: 92C7054341326A1F CRC64;

Query Match 67.4%; Score 62; DB 12; Length 93;  
Best Local Similarity 100.0%; Pred. No. 0.034;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTLGIVCPI 12  
|||||  
Db 82 LLMGTLGIVCPI 93

RESULT 12  
Q9QDH4 PRELIMINARY; PRT; 93 AA.  
AC Q9QDH4: 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE E7 PROTEIN (FRAGMENT).  
GN E7.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
CC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KE6-3;  
RA Lee H.P., Song Y.S., Kim J.W., Roh J.W., Park N.H., Kang S.B.;  
RT "The most commonly found Human Papillomavirus type 16 E6 and E7  
sequence variation tend to occur together."  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF187868; AAF13397.1; -  
DR InterPro: IPR000148; Papv1\_E7.  
DR Pfam: PF00527; E7; 1.  
FT NON\_TER 93  
SQ SEQUENCE 93 AA: 10452 MW: 83281BB2AE2C9A1F CRC64;

Query Match 67.4%; Score 62; DB 12; Length 93;  
Best Local Similarity 100.0%; Pred. No. 0.034;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTLGIVCPI 12  
|||||  
Db 82 LLMGTLGIVCPI 93

RESULT 13  
Q9QDH2 PRELIMINARY; PRT; 93 AA.  
AC Q9QDH2: 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE E7 PROTEIN (FRAGMENT).  
GN E7.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
CC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KE6-4;  
RA Lee H.P., Song Y.S., Kim J.W., Roh J.W., Park N.H., Kang S.B.;  
RT "The most commonly found Human Papillomavirus type 16 E6 and E7  
sequence variation tend to occur together."  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF187869; AAF13399.1; -  
DR InterPro: IPR000148; Papv1\_E7.  
DR Pfam: PF00527; E7; 1.  
FT NON\_TER 93  
SQ SEQUENCE 93 AA: 10452 MW: 83281BB2AE2C9A1F CRC64;

Query Match 67.4%; Score 62; DB 12; Length 93;  
Best Local Similarity 100.0%; Pred. No. 0.034;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLMGTLGIVCPI 12  
|||||  
Db 82 LLMGTLGIVCPI 93

RESULT 14

Job time: 252 sec

090C22  
 ID 090C22 PRELIMINARY; PRT; 98 AA.  
 AC 090C21:  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 58.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-E7-HK-1;  
 RA Chan P.K.S., Lam C.W., Chan M.Y.M., Li W.W.H., Cheung J.L.K.,  
 RA Cheng A.F.;  
 RT \*Sequence variation of human papillomavirus type 58 E7 among Chinese  
 RT with cancerous and noncancerous cervical lesions.\*;  
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF192269; AAF13690.1; -;  
 DR InterPro: IPR000148; PapvL\_E7;  
 DR Pfam: PF00527; E7; 1.  
 SQ SEQUENCE 98 AA: 10861 MW: 83657B91J355721 CRC64;

Query Match 64.1%; Score 59; DB 12; Length 98;  
 Best Local Similarity 68.8%; Pred. No. 0.1;  
 Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LMGTLGIVCIPICSOX 16  
 ||||| ||||| |||||  
 DB 83 LMGTCCTIVCPCSAQO 98

RESULT 15  
 ID 090C21 PRELIMINARY; PRT; 98 AA.  
 AC 090C21:  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 58.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-E7-HK-2;  
 RA Chan P.K.S., Lam C.W., Chan M.Y.M., Li W.W.H., Cheung J.L.K.,  
 RA Cheng A.F.;  
 RT \*Sequence variation of human papillomavirus type 58 E7 among Chinese  
 RT with cancerous and noncancerous cervical lesions.\*;  
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF192270; AAF13691.1; -;  
 DR InterPro: IPR000148; PapvL\_E7.  
 DR Pfam: PF00527; E7; 1.  
 SQ SEQUENCE 98 AA: 10976 MW: D70C8D4DDACE5D12 CRC64;

Query Match 64.1%; Score 59; DB 12; Length 98;  
 Best Local Similarity 68.8%; Pred. No. 0.1;  
 Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LMGTLGIVCIPICSOX 16  
 ||||| ||||| |||||  
 DB 83 LMGTCCTIVCPCSAQO 98

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 13, 2002, 09:53:46 ; Search time 66.5 seconds  
(without alignments)  
25.619 Million cell updates/sec

Title: US-09-664-225-66  
129  
Sequence: 1 KISEYRHVCYSLVGTLEQQYNK 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq\_1101:\*

- 1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT:\*
- 22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	23	AA895959	HPV 16 E6 protein
2	129	100.0	117	AA896050	Polyepitope polypeptide
3	129	100.0	158	AA822766	HPV E6 peptide
4	129	100.0	158	AA822766	Human papillomavirus
5	129	100.0	158	AA898420	Human papillomavirus
6	129	100.0	162	AAW35741	Human papillomavirus
7	129	100.0	188	AA863865	HPV16 E6/E7 protein
8	129	100.0	236	AA896052	Polyepitope polypeptide
9	129	100.0	243	AA899369	Papillomavirus E6/
10	129	100.0	263	AA827725	HPV 16 E6 protein
11	129	100.0	266	AA897561	Human papilloma vi

12	129	100.0	273	AA825376	HPV fusion protein
13	129	100.0	273	AA802632	Prot. D1/3-E6-His/H
14	129	100.0	292	AA825379	HPV fusion protein
15	129	100.0	292	AA802635	CLYRA-E6-His prote
16	129	100.0	371	AA825377	HPV fusion protein
17	129	100.0	371	AA802633	Prot. D1/3-E6-E7-H
18	129	100.0	390	AA825381	HPV fusion protein
19	129	100.0	390	AA802637	CLYRA-E6E7-His pro
20	126	97.7	162	AA835742	Human papillomavir
21	124	96.1	151	AA857808	HPV-16 E6 protein
22	119	92.2	29	AA830944	Peptide comprising
23	119	92.2	29	AA831018	Polyepitopic fragm
24	119	92.2	29	AA831110	A polyepitopic fra
25	116	89.9	172	AA897562	Human papilloma vi
26	101	78.3	20	AA815566	Human papilloma vi
27	99	76.7	172	AA897563	Immunopeptide #1 d
28	96	74.4	18	AA829574	Human papilloma vi
29	89	69.0	14	AA898447	Peptide GF59. Syn
30	87	67.4	14	AA840919	Human papillomavir
31	84	65.1	18	AA829575	Peptide GF60. Syn
32	79	61.2	149	AA898434	Human papillomavir
33	66	51.2	11	AA845456	Immunogenic peptid
34	65	50.4	150	AA898413	Human papillomavir
35	65	50.4	155	AA898452	Human papillomavir
36	63	48.8	10	AA895991	Human papillomavir
37	63	48.8	150	AA898395	Human papillomavir
38	63	48.8	150	AA898404	Human papillomavir
39	63	48.8	368	AAW01808	Human papillomavir
40	63	48.8	375	AAW01809	Human papillomavir
41	63	48.8	465	AAW01810	Human papillomavir
42	63	48.8	1587	AAW01816	Human papillomavir
43	62	48.1	10	AA895989	HPV 16 E6 A24 MHC-
44	61	47.3	10	AA845997	Immunogenic peptid
45	61	47.3	10	AA895974	HPV 16 E6 A1 MHC-D

## ALIGNMENTS

RESULT 1	
AA895959	standard; Peptide; 23 AA.
XX	
XX	AA895959;
AC	
XX	
DT	25-JUN-2001 (first entry)
XX	
DE	HPV 16 E6 protein fragment SEQ ID 66.
XX	
KW	Epitope; tumour antigen; antiviral; immunostimulatory; cervical cancer;
KW	human papillomavirus-associated disease; condyloma; cervical dysplasia;
KW	cervical dysplasia; major histocompatibility complex; MHC I.
XX	
OS	Human papillomavirus.
PN	WO200119408-A1.
XX	
PD	22-MAR-2001.
XX	
XX	
PE	18-SEP-2000; 2000WO-US25559.
XX	
PR	16-SEP-1999; 99US-0154665.
PR	16-SEP-1999; 99US-0398534.
PR	09-DEC-1999; 99US-0169846.
PR	09-DEC-1999; 99US-0458173.
XX	
PA	(ZYCO-) ZYCOS INC.
XX	
PI	Hedley ML, Urban RC, Chicz RM;
DR	WPI; 2001-265996/27.
XX	
PT	Novel nucleic acids encoding polypeptide polypeptides containing





PR 26-SEP-1990; 90US-0588384.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIB.  
 XX  
 PI Blake J, Chen L, Hellstrom I, Hellstrom K, Hu S L;  
 PI Thomas E K;  
 XX  
 DR WPI: 1992-132119/16.  
 XX  
 PT Immunogenic peptide(s) derived from E6 or E7 region of HPV16 -  
 PT and recombinant cells encoding them, useful in treatment and  
 PT prophylaxis of cervical warts or cancer resulting from HPV  
 PT infection  
 CC  
 PS Disclosure: Fig 7: 81pp; English.  
 XX  
 CC The peptide is the sequence of the human papillomavirus HPV 16 E6  
 CC nucleoprotein. Peptides corresponding to regions (pref. epitopic  
 CC regions) of HPV 16 E6 were synthesised by standard Merrifield  
 CC synthesis. Examples of such peptides are E6 1-20, 8-20, 119-134 or  
 CC 148-158. Compositions contg. these peptides, antihodies against the  
 CC peptides, or recombinant cells contg. the gene encoding the immuno-  
 CC genic peptides may be utilised in methods for inhibiting and treating  
 CC HPV infection and tumour initiation and progression e.g. in the  
 CC prevention or retardation of cervical warts and cervical carcinoma  
 CC resulting from HPV infection.  
 CC See also AAR22767.  
 CC  
 XX  
 SQ Sequence 158 AA:

Query Match 100.0%; Score 129; DB 13; Length 158;  
 Best local Similarity 100.0%; Pred. No. 1.1e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHYCYSLYGTLEQYNNK 23  
 ||||||||||||||||||||  
 DB 79 KISEYRHYCYSLYGTLEQYNNK 101

RESULT 4  
 AAY82462  
 ID AAY82462 standard; Protein; 158 AA.  
 XX  
 AC AAY82462;  
 XX  
 DT 30-JUN-2000 (first entry)  
 XX  
 DE Human papillomavirus E6 protein containing two zinc finger motifs.  
 XX  
 KW Chelated zinc finger; therapeutic; treatment; prophylaxis; HPV;  
 KW mammalian papillomavirus; antiviral; cytostatic; cervical cancer;  
 KW lesion; wart.  
 XX  
 OS Human papillomavirus.  
 XX  
 Key Location/Qualifiers  
 FT Misc-difference 37..73 /note= "forms a zinc finger motif"  
 FT Misc-difference 110..146 /note= "forms a zinc finger motif"  
 FT Misc-difference 110..146 /note= "forms a zinc finger motif"  
 PN WO200014063-A1.  
 XX  
 XX 16-MAR-2000.  
 XX  
 PF 03-SEP-1999; 99WO-AU00724.  
 XX  
 PR 04-SEP-1998; 98AU-0005733.  
 PR 15-JUL-1999; 99AU-0001645.  
 XX  
 PA (MOLE-) INST MOLECULAR & CELL BIOLOGY.  
 PA (HUGH-) HUGHES E J L.

XX  
 XX Bernard H, Tan YJ, Beerheide W, Ting AE, Sim MM;  
 XX  
 DR WPI: 2000-256917/22.  
 XX  
 PT Polysulfide and dithionodisulfide agents, useful for the treatment or  
 PT prophylaxis of diseases caused by mammalian papillomavirus, are  
 PT disruptors of a chelated metal cation domain in an MPV gene encoded  
 PT protein -  
 CC  
 PS Disclosure: Fig 1: 78pp; English.  
 XX  
 CC The present invention describes an agent used in the treatment or  
 CC prophylaxis of a disease caused or exacerbated by MPV (mammalian  
 CC papillomavirus) comprising a compound capable of reducing, inhibiting  
 CC or otherwise decreasing the activity of a protein encoded by an MPV  
 CC gene by facilitating disruption of a chelated metal cation domain  
 CC present in the protein. An agent of the present invention can be used  
 CC to treat cervical cancer or its HPV associated precursor lesions or  
 CC other HPV associated cancers and/or warts. The present sequence  
 CC represents a human papillomavirus E6 protein containing two zinc  
 CC finger motifs, as given in the exemplification of the present  
 CC invention.  
 CC  
 XX  
 SQ Sequence 158 AA:

Query Match 100.0%; Score 129; DB 21; Length 158;  
 Best local Similarity 100.0%; Pred. No. 1.1e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHYCYSLYGTLEQYNNK 23  
 ||||||||||||||||||||  
 DB 79 KISEYRHYCYSLYGTLEQYNNK 101

RESULT 5  
 AAB98420  
 ID AAB98420 standard; Protein; 158 AA.  
 XX  
 AC AAB98420;  
 XX  
 DT 22-AUG-2001 (first entry)  
 XX  
 DE Human papillomavirus protein HPV16 E6.  
 XX  
 KW Human papillomavirus; human leukocyte antigen; HLA; immune response;  
 KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;  
 KW neoplastic growth; antiviral.  
 XX  
 OS Human papillomavirus.  
 XX  
 PN WO200141799-A1.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 11-DEC-2000; 2000WO-US33549.  
 XX  
 PR 10-DEC-1999; 99US-0172705.  
 PR 15-AUG-2000; 2000US-0641528.  
 XX  
 PA (EPTM-) EPTIMUNE INC.  
 XX  
 PI Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;  
 PI WPI: 2001-381497/40.  
 DR  
 XX  
 PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for  
 PT treating HPV infections -  
 XX  
 PS Disclosure: Page 20-21; 756pp; English.  
 XX  
 CC The present invention describes an isolated prepared human papillomavirus

CC (HPV) epitope (1). (1) has antiviral activity, and can be used in  
 CC vaccine production. Peptides and corresponding nucleic acid compositions  
 CC from the present invention are useful for stimulating an immune response  
 CC to HPV by stimulating the production of CTL or HTL responses,  
 CC specifically in the treatment or prophylaxis of HPV infection, in persons  
 CC who have not manifested symptoms e.g. genital warts or neoplastic growth.  
 CC The peptides can also be used in a tetramer staining assay to assess  
 CC peripheral blood mononuclear cells for the presence of antigen-specific  
 CC CTLs following exposure to a pathogen or immunogen, and as reagents to  
 CC evaluate immune recall responses or evaluate the efficacy of a vaccine.  
 CC The vaccine compositions are useful for removing warts or treating HPV  
 CC infections. The epitopes for inclusion in an epitope-base vaccine may  
 CC be selected from conserved regions of viral or tumour-associated  
 CC antigens, which reduces the likelihood of escape mutants, also  
 CC immunosuppressive epitopes that may be present in whole antigens can be  
 CC avoided with the use of epitope-base vaccines. An additional advantage  
 CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
 CC the composition of the epitopes achieving enhanced immunogenicity, the  
 CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
 CC to AAB98477 represent polypeptide sequences used in the exemplification  
 CC of the present invention.

SO Sequence 158 AA;

Query Match 100.0%; Score 129; DB 22; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTLEEOYRK 23  
 Db 79 Kiseyryhvcyslygtleegynk 101

RESULT 6

AAM35741  
 ID AAM35741 standard; Protein: 162 AA.

XX AAM35741;

DT 16-FEB-1998 (first entry)

XX Human papillomavirus type 16 E6 protein.

XX Complete genome; circular; human papillomavirus type 16; HPV16 E6;  
 XX cervical dysplasia; cervical cancer; cervical smear.

OS Human papillomavirus type 16.

XX Key Location/Qualifiers

FT Misc-difference 90

FT /note- "Mutated to Val in the variant"

FT Misc-difference 159  
 FT /note- "End of protein sequence even though 3 amino  
 FT acid residues are given following on"

XX US5679509-A.

PD 21-OCT-1997.

XX 28-SEP-1993; 93US-0127906.

XX 30-SEP-1994; 94US-0316239.

XX 28-SEP-1993; 93US-0127906.

PA (UYNE-) UNIV NEW MEXICO STATE.

PI Parmenter CA, Wheeler CM;

DR WPI: 1997-525714/48.

XX N-PSDB; AAT94723.

PT Evaluating risk of cervical dysplasia or cervical cancer - by

PT detecting variant form of human papilloma virus 16  
 XX Claim 7; Column 23-24; 33pp; English.

CC Methods have been developed for distinguishing a subset of human  
 CC papilloma virus (HPV) that is associated with an increased risk of  
 CC developing cervical dysplasia or cervical cancer. The methods involve:  
 CC (1) preparing a cervical sample to expose any HPV-16 E6 gene in the  
 CC sample and determining if the base at position 350 of the E6 gene  
 CC (see AAT94723 and AAT94724 for comparison) is T or G, where the presence  
 CC of G at position 350 is associated with an increased risk of developing  
 CC cervical dysplasia or cervical cancer; and (2) preparing a cervical  
 CC sample to expose any HPV-16 E6 protein in the sample and determining  
 CC if the amino acid at position 83 of the protein (see position 90 in  
 CC AAM35741 and AAM35742 for comparison) is Val or Leu, where the presence  
 CC of Val at position 83 that is associated with an increased risk of  
 CC developing cervical dysplasia or cervical cancer. The present sequence  
 CC represents the reference protein sequence for HPV-16 E6. The 350G  
 CC variant correlates well with Pap scores: 350T:350G ratios among 45  
 CC HPV16 samples were 10:4 for negative Pap scores; 4:2 for CIN I, 1:6 for  
 CC CIN II; 2:9 for CIN III; 0:3 for cancer.

SO Sequence 162 AA;

Query Match 100.0%; Score 129; DB 18; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTLEEOYRK 23  
 Db 79 Kiseyryhvcyslygtleegynk 101

RESULT 7

AAR63865  
 ID AAR63865 standard; Protein: 188 AA.

XX AAR63865;

DT 28-JUN-1995 (first entry)

XX HPV16 E6/E7 proteins.

XX HPV; HPV16; E6 protein; E7 protein; diagnosis; cervical dysplasia;  
 XX cervix cancer.

OS Human papillomavirus strain 16.

XX Key Location/Qualifiers

FT Protein 1..158

FT /label- E6-protein

FT Protein 159..188  
 FT /label- E7-protein

XX W09426934-A.

PD 24-NOV-1994.

XX 06-MAY-1994; 94MO-US05085.

XX 06-MAY-1993; 93US-0058920.

PA (BAXT) BAXTER DIAGNOSTICS INC.

PI Brown JT;

DR WPI: 1995-006821/01.

XX P-PSDB; AAO75470.

PT Human papilloma virus detection assay - by amplification using  
 PT self sustained sequence replication and hybridisation with a  
 PT detector probe

XX PS Disclosure; Page 24-26; 79pp; English.  
XX CC The sequences of the E6 and E7 polypeptide-encoding regions of human  
CC papillomavirus (HPV) 16 and 18 are given in AA075470-71 and the  
CC encoded proteins in AAK63865-66, respectively. Probes and primers  
CC based on these sequences were used for HPV infection diagnosis;  
CC expression of E6 and E7 is diagnostic for cervical cancer or pre-  
CC malignant states.  
XX CC  
SQ Sequence 188 AA;  
  
Query Match 100.0%; Score 129; DB 16; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1,4e-11;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 KISEYRHVCYSLYGTLEQYNK 23  
Db 79 KISEYRHVCYSLYGTLEQYNK 101  
|||||  
RESULT 8  
ID AAB96052 standard; Peptide; 236 AA.  
XX AAB96052;  
AC  
XX 25-JUN-2001 (first entry)  
XX  
DE Polypeptide polypeptide SEQ ID 157.  
XX  
XX Epitope: tumour antigen; antiviral; immunostimulatory; cervical cancer;  
KM human papillomavirus-associated disease; condyloma; cervical dysplasia;  
KM cervical dysplasia; major histocompatibility complex; MHC I.  
XX  
XX Synthetic.  
OS  
XX  
XX WO200119408-A1.  
PN  
XX  
XX 22-MAR-2001.  
PD  
XX  
XX 18-SEP-2000; 2000WO-US25559.  
PF  
XX  
XX 16-SEP-1999; 99US-0154665.  
PR 16-SEP-1999; 99US-0398534.  
PR 09-DEC-1999; 99US-0169846.  
PR 09-DEC-1999; 99US-0458173.  
PR  
XX  
XX (2YCO-) 2YCO5 INC.  
PA  
XX  
XX Hedley ML, Urban RC, Chicx RM.  
PI  
XX  
XX WPI; 2001-265996/27.  
DR  
XX  
XX Novel nucleic acids encoding polypeptide polypeptides containing  
PT multiple epitopes from one or more proteins, useful for treating tumors  
PT and as vaccines against pathogenic agents  
XX  
XX Disclosure; Fig 5; 64pp; English.  
PS  
XX  
XX This invention relates to polynucleotides encoding a hybrid polypeptide  
CC comprising a signal sequence and three segments that are either  
CC contiguous or separated by a spacer amino acid or spacer peptide. The  
CC invention specifically details polynucleotides encoding a polypeptide  
CC peptide where the peptide segments are tumour antigens or a naturally  
CC occurring protein of a pathogenic agent. The polypeptide peptides exhibit  
CC antiviral and immunostimulatory activity. The polynucleotide and  
CC polypeptide peptides are useful for eliciting an immune response in a  
CC mammal. The polynucleotide and protein are useful as vaccines for  
CC treating tumours and pathogenic infections. The polynucleotide is also  
CC useful for preventing or treating human papillomavirus (HPV)-associated  
CC diseases, particularly exophytic condyloma, flat condyloma, cervical

CC cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV  
CC infection, cervical dysplasia, high grade squamous intraepithelial  
CC lesions, and anal HPV infection. The polynucleotide and polypeptide are  
CC useful for generating or enhancing prophylactic or therapeutic immune  
CC response against pathogens, tumours or autoimmune diseases in a  
CC population of individuals having diverse MHC allotypes, as positive  
CC controls in T cell stimulation assays in vitro, and as tools to  
CC understand processing of epitopes within cells. Peptides  
CC AAB95894 - AAB96037 and AAB96044 - AAB96048 represent major  
CC histocompatibility complex 1 (MHC I) associated tumour and pathogen  
CC antigens. The peptides can be used as part of the polypeptide proteins of  
CC the invention. Also included are examples of the polypeptide proteins  
CC represented by AAB96050 - AAB96052, and localisation signal peptides  
CC AAB96038 - AAB96043 and AAB96049 which can be used in the construction of  
CC the polypeptide peptides.  
XX  
SQ Sequence 236 AA;  
  
Query Match 100.0%; Score 129; DB 22; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1,8e-11;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 KISEYRHVCYSLYGTLEQYNK 23  
Db 45 KISEYRHVCYSLYGTLEQYNK 67  
|||||  
RESULT 9  
ID AAW99369 standard; peptide; 243 AA.  
XX AAW99369;  
AC  
XX 21-MAY-1999 (first entry)  
XX  
DE Papillomavirus E6/protein F signal peptide fusion protein.  
XX  
XX Antitumour; immunogen; intracellular localisation; cancer; dysplasia;  
KM signal peptide; papillomavirus; infection; cervix uteri.  
KM  
XX  
XX Chimeric - Human papillomavirus.  
OS  
XX  
XX WO9903885-A1.  
PN  
XX  
XX 28-JAN-1999.  
PD  
XX  
XX 17-JUL-1998; 98WO-FR01576.  
PF  
XX  
XX 18-JUL-1997; 97FR-0009152.  
PR  
XX  
XX (TRGE ) TRANSGENE SA.  
PA  
XX  
XX Balloul JM, Bizouarne N, Kieny MP.  
PI  
XX  
XX WPI; 1999-132161/11.  
DR  
XX  
XX Antitumour composition containing immunogenic polypeptide with  
PT altered localisation - or vector expressing this polypeptide,  
PT particularly for treating or preventing cervical cancer associated  
PT with human papilloma virus  
XX  
PS Claim 16; Page 54-55; 74pp; French.  
XX  
XX The invention relates to an antitumour composition containing a  
CC therapeutic or prophylactic agent or one or more immunogenic  
CC polypeptides, where at least one polypeptide is modified so that  
CC its intracellular localisation is different from that of the native  
CC polypeptide. This sequence represents an example of a modified  
CC immunogenic protein and corresponds to the E6 protein from human  
CC papillomavirus fused to the measles virus protein F signal peptide.  
CC The compositions, vectors and particles are used to treat or prevent

CC cancers and tumours, specifically those associated with papillomavirus  
 CC infection (e.g. cancer or low grade dysplasia of the cervix uteri).  
 XX  
 SO Sequence 243 AA;

Query Match 100.0%; Score 129; DB 20; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 1,8e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTLEQOYNK 23  
 ||||||||||||||||||  
 DB 107 KISEYRHVCYSLYGTLEQYNK 129

RESULT 10  
 AAR27725  
 ID AAR27725 standard; Protein: 263 AA.

AC AAR27725;  
 DT 09-MAR-1993 (first entry)  
 XX

DE HPV 16 E6 protein fragment.

XX  
 KW Virus vector; vaccinia virus; papillomavirus; HPV; human;  
 amplification; immunotherapeutic.

OS Human papillomavirus 16.

XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..159  
 PT /note= "HPV-16 E6 protein"

PN W0921636-A.

XX 01-OCT-1992.

PD 10-MAR-1992; 92WO-GB00424.

XX 14-MAR-1991; 91GB-0005383.

XX (IMMU ) IMMUNOLOGY LTD.

PI Bourneill MEG, Inglis SC, Munro AJ;

DR WPI: 1992-349219/42.

XX N-PSDB; AAQ29389.

PT Recombinant virus vectors encoding human papillomavirus proteins  
 - for treating and vaccinating against HPV infections and  
 PT conditions caused by them, such as cervical cancer

PS Disclosure; Fig 1a; 83pp; English.

XX The fragment of DNA contg. the HPV-16 E6/E7 coding region was  
 CC prep'd. by PCR from plasmid pAR322/HPV16 (Durst et al., PNAS, 80:  
 CC 3812 (1983)) using oligonucleotides S05 and S06. The prod. of the  
 CC third reading frame is the HPV-16 E6 protein whereas the second  
 CC reading frame encodes HPV-16 E7. The E6 and E7 ORFs are fused  
 CC together to form a single continuous ORF via site directed mutagenesis  
 CC and the immortalising potential of E7 is removed by altering two key  
 CC codons of the HPV E7 sequence. The single ORF of HPV-16 E6/E7 may be  
 CC inserted into vaccinia virus DNA at neutral sites (pref. by inserting  
 CC two sets of the DNA in opposite orientations to overcome the problem  
 CC of intertypic recombination) to make a recombinant virus vector for  
 CC use immunotherapeutically to activate cells of the immune system  
 CC against HPV. See also AAR27725-43.

XX Sequence 263 AA;

Query Match 100.0%; Score 129; DB 13; Length 263;

Best Local Similarity 100.0%; Pred. No. 2e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTLEQOYNK 23  
 ||||||||||||||||||  
 DB 80 KISEYRHVCYSLYGTLEQYNK 102

RESULT 11  
 AAR97561  
 ID AAR97561 standard; Protein: 266 AA.

AC AAR97561;

DT 11-JAN-1997 (first entry)

DE Human papilloma virus E6/E7 protein variant.

XX  
 KW Human papilloma virus; E6; E7; deletion mutant; HPV;  
 KW immune response; humoral immune response; cellular immune response;  
 KW vaccine.

OS Human papilloma virus.

XX W09619496-A1.

PD 27-JUN-1996.

XX 20-DEC-1995; 95WO-A000868.

XX 20-DEC-1994; 94AU-0000157.

XX (CSIC-) CSL LTD.

PA (UYOU ) UNIV QUEENSLAND.

PI Cox J, Edwards SJ, Frazer I, Webb EA;

DR WPI: 1996-309518/31.

XX N-PSDB: AAT31833.

PT Vaccine variants of human papilloma virus antigens - contain  
 PT variants of E6 and/or E7 protein, pref. deletion mutants, and are  
 PT used to treat or prevent HPV infection

PS Example 1; Page 15-16; 37pp; English.

XX A variant of the human papilloma virus (HPV) E6 or E7 protein which  
 CC elicits a humoral and/or cellular immune response against HPV can be  
 CC used in vaccines against HPV or to treat HPV infection. The variant  
 CC is preferably a deletion mutant comprising at least half, and  
 CC preferably two-thirds of full length E6 or E7 protein starting from  
 CC the N- or C-terminal, or is a full length E6 moiety fused to a full  
 CC length E7 moiety. The variant optionally has a linkage moiety and a  
 CC foreign protein or peptide which facilitates the purification of,  
 CC and enhances the immunogenicity of, the fusion protein. This  
 CC sequence is a full length E6/E7 fusion protein.

XX Sequence 266 AA;

Query Match 100.0%; Score 129; DB 17; Length 266;  
 Best Local Similarity 100.0%; Pred. No. 2e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTLEQOYNK 23  
 ||||||||||||||||||  
 DB 79 KISEYRHVCYSLYGTLEQYNK 101

RESULT 12  
 AAY25376  
 ID AAY25376 standard; Protein: 273 AA.

```

AC  AAY25376;
XX
DT  06-SEP-1999 (first entry)
DE  HPV fusion protein D1/3-E6-His/HPV16.
XX
KM  Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KW  immunological fusion partner; CPG oligonucleotide; immune response;
XX  HPV antigen; prevention; treatment.
OS  Synthetic.
XX  Human Papillomavirus.
XX  WO933868-A2.
XX
PD  08-JUL-1999.
XX
PF  18-DEC-1998; 98WO-EP08563.
XX
PR  24-DEC-1997; 97GB-0027262.
XX
PA  (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI  Dalemans WJ, Gerard CMG;
XX
DR  WPI; 1999-405485/34.
XX  N-PSDB; AAX78792.
XX
PT  Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PS  induce immune response to HPV
XX  Example II; Page 48-49; 62pp; English.
XX
CC  AAX78791-X78801 represent nucleic acid sequences which encode novel
CC  constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC  HPV (represented in AAY25375-Y25386). These constructs are optionally
CC  linked to an immunological fusion partner and an immunomodulatory CPG
CC  oligonucleotide. The products of the invention can be used to induce an
CC  immune response in a patient to an HPV antigen. They can also be used
CC  for preventing or treating HPV induced tumours.
XX
SQ  Sequence 273 AA;

Query Match      100.0%; Score 129; DB 20; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  1 KISEYRHYCYSLYGTTLEQYNK 23
    |||||||||||||||||||
Db  185 kiseyrhycyslygttleqgynk 207

RESULT 13
AAY02632
ID  AAY02632 standard; Protein; 273 AA.
XX
AC  AAY02632;
XX
DT  22-JUN-1999 (first entry)
XX
DE  Prot.D1/3-E6-His/HPV16 protein.
XX
KM  Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
KW  tumour; lesion; benign; malignant; virus; infection.
XX
OS  Chimeric - Human papillomavirus.
XX  Chimeric - Haemophilus influenzae.
XX
PN  WO910375-A2.
XX
PD  04-MAR-1999.
XX

```

```

PF  17-AUG-1998; 98WO-EP05285.
XX
PR  22-AUG-1997; 97GB-0017953.
XX
PA  (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI  Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
XX  Lombardo-Bencheikh A;
XX
DR  WPI; 1999-190587/16.
XX  N-PSDB; AAX29781.
XX
PT  Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
PS  treatment or prophylaxis of HPV induced lesions
XX
XX  Disclosure: Fig 3; 95pp; English.
XX
CC  This sequence represents a chimeric E6 or E7 protein or E6/E7 fusion
CC  protein from Human papillomavirus (HPV) linked to an immunological fusion
CC  partner. In this case, a fragment of the Haemophilus influenzae B
CC  protein D. The sequence also contains a histidine tag at the C-terminus
CC  of the encoded protein. The protein can be used in a vaccine, for
CC  immuno-therapeutically treating HPV induced tumour lesions (benign or
CC  malignant) and preventing HPV viral infection.
XX
SQ  Sequence 273 AA;

Query Match      100.0%; Score 129; DB 20; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  1 KISEYRHYCYSLYGTTLEQYNK 23
    |||||||||||||||||||
Db  185 kiseyrhycyslygttleqgynk 207

RESULT 14
AAY25379
ID  AAY25379 standard; Protein; 292 AA.
XX
AC  AAY25379;
XX
DT  06-SEP-1999 (first entry)
XX
DE  HPV fusion protein CLYTA-E6-His/HPV16.
XX
KM  Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KW  immunological fusion partner; CPG oligonucleotide; immune response;
XX  HPV antigen; prevention; treatment.
XX
OS  Synthetic.
XX  Human papillomavirus.
XX  WO933868-A2.
XX
PD  08-JUL-1999.
XX
PF  18-DEC-1998; 98WO-EP08563.
XX
PR  24-DEC-1997; 97GB-0027262.
XX
PA  (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI  Dalemans WJ, Gerard CMG;
XX
DR  WPI; 1999-405485/34.
XX  N-PSDB; AAX78795.
XX
PT  Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PS  induce immune response to HPV
XX  Example VI; Page 52-53; 62pp; English.
XX

```

XX AAY78791-X78801 represent nucleic acid sequences which encode novel  
 CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from  
 CC HPV (represented in AAY25375-Y25386). These constructs are optionally  
 CC linked to an immunological fusion partner and an immunomodulatory CpG  
 CC oligonucleotide. The products of the invention can be used to induce an  
 CC immune response in a patient to an HPV antigen. They can also be used  
 CC for preventing or treating HPV induced tumours.  
 XX  
 SQ Sequence 292 AA;

Query Match 100.0%; Score 129; DB 20; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTTLEQGYNK 23  
 ||||||||||||||||||  
 DB 204 kiseyrhvcyslygttleqgynk 226

## RESULT 15

AAV02635  
 ID AAY02635 standard; Protein: 292 AA.  
 XX  
 AC AAY02635;  
 XX  
 DT 22-JUN-1999 (first entry)  
 XX  
 DE CLYTA-E6-His protein.  
 XX  
 KW Chimeric; E6; E7; fusion protein; CLYTA; vaccine; immunotherapy;  
 KW tumour; lesion; benign; malignant; virus; infection.  
 XX  
 OS Chimeric - Human Papillomavirus.  
 OS Chimeric - Streptococcus pneumoniae.  
 XX  
 PN W09910375-A2.  
 XX  
 XX 04-MAR-1999.  
 PD  
 XX 17-AUG-1998; 98WO-EP05285.  
 PF  
 XX 22-AUG-1997; 97GB-0017953.  
 PR  
 (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PA Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;  
 XX Lombardo-Bencheikh A;  
 PI  
 XX WPI: 1999-190587/16.  
 DR N-PSDB; AAX29784.  
 DR  
 PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for  
 PT treatment or prophylaxis of HPV induced lesions  
 XX  
 PS Disclosure: Fig 10; 95pp; English.  
 XX  
 CC This sequence represents a chimeric E6 or E7 protein or E6/E7 fusion  
 CC protein from Human papillomavirus (HPV) linked to an immunological fusion  
 CC partner, in this case, a fragment of the Streptococcus pneumoniae CLYTA  
 CC protein of the encoded protein. The protein can be used in a vaccine,  
 CC for immuno-therapeutically treating HPV induced tumour lesions (benign  
 CC or malignant) and preventing HPV viral infection.  
 XX  
 SQ Sequence 292 AA;

DB 204 kiseyrhvcyslygttleqgynk 226  
 ||||||||||||||||||

Search completed: February 13, 2002, 09:59:18  
 Job time: 332 sec

Query Match 100.0%; Score 129; DB 20; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTTLEQGYNK 23

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 09:54:11 ; Search time 32.4 Seconds  
(without alignments)  
15.975 Million cell updates/sec

Title: US-09-664-225-66  
Perfect score: 129  
Sequence: 1 KISEYRHYCYSLYGTLEQYKN 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents-AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTOTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	162	1	US-08-316-239B-3	Sequence 3, Appli
2	129	100.0	172	3	US-08-860-165-12	Sequence 12, Appl
3	129	100.0	172	4	US-09-359-382-12	Sequence 12, Appl
4	129	100.0	266	4	US-08-860-165-10	Sequence 10, Appl
5	129	100.0	266	4	US-09-359-382-10	Sequence 10, Appl
6	126	97.7	162	1	US-08-316-239B-4	Sequence 4, Appli
7	114	88.4	182	1	US-08-117-083-10	Sequence 10, Appl
8	101	78.3	20	2	US-08-934-915-44	Sequence 44, Appl
9	101	78.3	20	2	US-08-934-915-163	Sequence 163, App
10	99	76.7	172	4	US-08-860-165-14	Sequence 14, Appl
11	99	76.7	172	4	US-09-359-382-14	Sequence 14, Appl
12	66	51.2	11	3	US-08-159-339A-1171	Sequence 1171, Ap
13	59	45.7	158	2	US-08-247-904B-10	Sequence 10, Appl
14	59	45.7	158	3	US-08-767-942A-19	Sequence 19, Appl
15	59	45.7	271	1	US-08-117-083-14	Sequence 14, Appl
16	57	44.2	32	1	US-08-466-285-4	Sequence 4, Appli
17	56	43.4	9	3	US-08-159-339A-76	Sequence 76, Appl
18	53	41.1	10	3	US-08-159-339A-77	Sequence 77, Appl
19	53	41.1	20	2	US-08-934-915-45	Sequence 45, Appl
20	53	41.1	20	2	US-08-934-915-164	Sequence 164, App
21	51	39.5	9	3	US-08-159-339A-563	Sequence 563, App
22	48	37.2	9	3	US-08-159-339A-235	Sequence 235, App
23	48	37.2	9	3	US-08-159-339A-251	Sequence 251, App
24	47	36.4	9	3	US-08-159-339A-78	Sequence 78, Appl
25	47	36.4	9	3	US-08-159-339A-134	Sequence 134, App
26	47	36.4	9	3	US-08-159-339A-247	Sequence 247, App
27	46	35.7	1138	1	US-07-973-320-2	Sequence 2, Appli

28	46	35.7	1138	1	US-07-973-320-4	Sequence 4, Appli
29	46	35.7	1167	1	US-08-100-709-2	Sequence 2, Appli
30	46	35.7	1167	1	US-08-176-865-2	Sequence 2, Appli
31	46	35.7	1167	1	US-08-474-038-2	Sequence 2, Appli
32	46	35.7	1167	2	US-08-779-046-2	Sequence 2, Appli
33	46	35.7	1167	2	US-08-881-340-2	Sequence 2, Appli
34	45	34.9	9	3	US-08-881-340-2	Sequence 230, App
35	45	34.9	9	3	US-08-159-339A-230	Sequence 230, App
36	45	34.9	10	3	US-08-159-339A-234	Sequence 75, Appl
37	44	34.1	1168	1	US-08-448-170-6	Sequence 6, Appli
38	44	34.1	1168	1	US-08-961-803-8	Sequence 8, Appli
39	43.5	33.7	174	2	US-08-401-530A-6	Sequence 6, Appli
40	43.5	33.7	174	2	US-08-709-662-6	Sequence 6, Appli
41	43.5	33.7	500	4	US-08-158-735A-2	Sequence 2, Appli
42	43.5	33.7	532	2	US-08-481-337A-6	Sequence 6, Appli
43	43.5	33.7	532	4	US-09-382-256-6	Sequence 6, Appli
44	43.5	33.7	532	4	US-09-382-256-14	Sequence 14, Appli
45	43.5	33.7	532	4	US-08-158-735A-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-08-316-239B-3  
; Sequence 3, Application US/08316239B  
; Patent No. 5679509  
; GENERAL INFORMATION:  
; APPLICANT: Wheeler, Cosette M.  
; APPLICANT: Parmenter, Cheryl A.  
; TITLE OF INVENTION: Methods and a Diagnostic Aid for  
; TITLE OF INVENTION: Distinguishing a Subset of HPV that is Associated with an  
; TITLE OF INVENTION: Increased Risk of Developing Cervical Dysplasia and  
; NUMBER OF INVENTIONS: 4  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Jagtiani & Associates  
; STREET: 6126 Rocky Way Court  
; CITY: Centreville  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 20120-3400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/316,239B  
; FILING DATE: 30-SEP-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jagtiani, Ajay A.  
; REGISTRATION NUMBER: 35,205  
; REFERENCE/DOCKET NUMBER: UNME-0001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 817-9453  
; TELEFAX: (703) 803-9387  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 162 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; US-08-316-239B-3  
Query Match 100.0%, Score 129, DB 1, Length 162;  
Best Local Similarity 100.0%, Pred. NO. 2.6e-12;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 KISEYRHYCYSLYGTTLEQQYNK 23
         |||||
Db      79 KISEYRHYCYSLYGTTLEQQYNK 10

```

```

RESULT      2
US-08-860-165-12
: Sequence 12. Application US/08860165A
Patent No. 6004557
GENERAL INFORMATION:
APPLICANT: EDWARDS, Stirling John
APPLICANT: COX, John Cooper
APPLICANT: WEBB, Elizabeth Ann
APPLICANT: FRAZER, Ian
TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
FILE REFERENCE: 17227/130
CURRENT APPLICATION NUMBER: US/08/860,165A
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: PCT/AU95/00868
EARLIER FILING DATE: 1995-12-20
EARLIER APPLICATION NUMBER: AU PN0157
EARLIER FILING DATE: 1994-12-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 12
LENGTH: 172
TYPE: PRY
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-12

```

Query Match	100.0%;	Score 129;	DB 3;	length 172;
Best Local Similarity	100.0%;	Pred. No. 2,8e-12;		
Matches	23;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0
OY	1	KISEYRHHYCYSLXGTTLEDOYNNK	23	
db	17	KISEYRHHYCYSLXGTTLEDOYNNK	39	

```

RESULT      3
US-09-359-382-12
: Sequence 12, Application US/09359382
: Patent No. 6306397
:
: GENERAL INFORMATION:
:   APPLICANT: EDWARDS, Stirling John
:   APPLICANT: COX, John Cooper
:   APPLICANT: WEBB, Elizabeth Ann
:   APPLICANT: PRATER, Ian
:
: TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENNS
:
: FILE REFERENCE: 017227/0148
:
: CURRENT APPLICATION NUMBER: US/09/359, 382
:
: CURRENT FILING DATE: 1999-07-23
:
: EARLIER APPLICATION NUMBER: US 08/860,165
:
: EARLIER FILING DATE: 1997-09-22
:
: EARLIER APPLICATION NUMBER: PCT/AU95/00868
:
: EARLIER FILING DATE: 1995-12-20
:
: EARLIER APPLICATION NUMBER: AU PNO157/94
:
: EARLIER FILING DATE: 1994-12-20
:
: NUMBER OF SEQ ID NOS: 27
:
: SOFTWARE: patentlun Ver. 2.0
:
: SEQ ID NO 12
:
: LENGTH: 172
:
: TYPE: PRP
:
: ORGANISM: Human papillomavirus type 16
:
: US-09-359-382-12

```

Query Match	100.0%;	Score 129;	DB 4;	length 172;
Fast Local Similarity	100.0%;	Pred. No. 2.8e-12;		
Matches	23;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0

QY 1 KISEYRHYCYSLYGTTLEQYNK 23  
|||  
Db 17 KISEYRHYCYSLYGTTLEQYNK 39

```

RESULT      4
US-08-660-165-10
; Sequence 10. Application US/08860165A
; Patent No. 6004557
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 17227/130
; CURRENT APPLICATION NUMBER: US/08/860.165A
; CURRENT FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU P90157
; EARLIER FILING DATE: 1994-12-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 266
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-10

```

Query Match	100.0%	Score 129	DB 3	Length 266
Best Local Similarity	100.0%	Pred. No. 4	66-13	
Matches 23	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	KISEYRHHYCYSLYGTLLREQYNNK	23	
Db	79	KISEYRHHYCYSLYGTLLREQYNNK	101	

```

RESULT      5
US-09-359-382-10
? Sequence 10, Application US/09355382
? Patent No. 6506397
? GENERAL INFORMATION:
? APPLICANT: EDMARDS, Stirling John
? APPLICANT: COX, John Cooper
? APPLICANT: WEBB, Elizabeth Ann
? APPLICANT: PRATER, Ian
? TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENNS
? FILE REFERENCE: 01727/0148
? CURRENT APPLICATION NUMBER: US/09/359, 382
? CURRENT FILING DATE: 1999-07-23
? EARLIER APPLICATION NUMBER: US 08/860, 165
? EARLIER FILING DATE: 1997-09-22
? EARLIER APPLICATION NUMBER: PCT/AU95/00868
? EARLIER FILING DATE: 1995-12-20
? EARLIER APPLICATION NUMBER: AU PNO157/94 .
? EARLIER FILING DATE: 1994-12-20
? NUMBER OF SEQ ID NOS: 27
? SOFTWARE: patentln Ver. 2.0
? SEQ ID NO 10
? LENGTH: 266
? TYPE: PRT
? ORGANISM: Human papillomavirus type 16
US-09-359-382-10

```

Query Match	100.0%;	Score 129;	DB 4;	Length 266;
Best Local Similarity	100.0%;	Pred. No. 4.6e-12;		



Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCSLYGTTLEOQYNK 23  
|||||  
Db 79 KISEYRHVCSLYGTTLEOQYNK 101

RESULT 6  
US-08-316-239B-4  
; Sequence 4, Application US/08316239B  
; Patent No. 5679509  
; GENERAL INFORMATION:  
; APPLICANT: Wheeler, Cosette M.  
; APPLICANT: Parmenter, Cheryl A.  
; TITLE OF INVENTION: Methods and a Diagnostic Aid for  
; TITLE OF INVENTION: Distinguishing a Subset of HPV that is Associated with an  
; TITLE OF INVENTION: Increased Risk of Developing Cervical Dysplasia and  
; TITLE OF INVENTION: Cervical Cancer  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jagtiani & Associates  
; STREET: 6126 Rocky Way Court  
; CITY: Centreville  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 20120-3400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/316,239B  
; FILING DATE: 30-SEP-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jagtiani, Ajay A.  
; REGISTRATION NUMBER: 35,205  
; REFERENCE/DOCKET NUMBER: UNME-0001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 817-9453  
; TELEFAX: (703) 803-9387  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 162 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; US-08-316-239B-4

Query Match 97.7%; Score 126; DB 1; Length 162;  
Best Local Similarity 95.7%; Pred. No. 7,4e-12;  
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCSLYGTTLEOQYNK 23  
|||||  
Db 79 KISEYRHVCSLYGTTLEOQYNK 101

RESULT 7  
US-08-117-083-10  
; Sequence 10, Application US/08117083  
; Patent No. 5719054  
; GENERAL INFORMATION:  
; APPLICANT: Bourisnell, Michael E.  
; APPLICANT: Ingalls, Stephen C.  
; APPLICANT: Munro, Alan J.  
; TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human  
; TITLE OF INVENTION: Papilloma Virus Proteins  
; NUMBER OF SEQUENCES: 70

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Walter H. Dreger  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/117,083  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Walter H.  
; REGISTRATION NUMBER: 24,190  
; REFERENCE/DOCKET NUMBER: A-58783  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 182 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: protein  
; LOCATION: 1..182  
; OTHER INFORMATION: /note="Xaa refers to stop codon in  
; OTHER INFORMATION: the open reading frame."  
US-08-117-083-10

Query Match 88.4%; Score 114; DB 1; Length 182;  
Best Local Similarity 90.9%; Pred. No. 5,4e-10;  
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ISEYRHVCSLYGTTLEOQYNK 23  
|||||  
Db 61 ISEYRHVCSLYGTTLEOQYEE 82

RESULT 8  
US-08-934-915-44  
; Sequence 44, Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; APPLICANT: DILLNER, LENA  
; APPLICANT: CHENG, HWE-MING  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
; TITLE OF INVENTION: USEFUL IN IMMUNOSSAY FOR  
; NUMBER OF SEQUENCES: 193  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MASON & ASSOCIATES, P.A.  
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
; CITY: CLEARWATER  
; STATE: FLORIDA  
; COUNTRY: U.S.A.  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 3.0  
; SOFTWARE: Microsoft word 6.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/934,915  
FILING DATE: 22-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/949,836  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LOUISE A. FOUTCH  
REGISTRATION NUMBER: 37,133  
REFERENCE/DOCKET NUMBER: 1946,6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 813-538-3800  
TELEFAX: 813-538-3820  
TELEX:  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-934-915-44

Query Match 78.3%; Score 101; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 4.1e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCSLYGTTLE 18  
Db 3 KISEYRHVCSLYGTTLE 20

RESULT 9  
US-08-934-915-163  
Sequence 163 Application US/08934915  
Patent No. 5932812  
GENERAL INFORMATION:  
APPLICANT: DILLNER, JOAKIM  
APPLICANT: DILLNER, LENA  
APPLICANT: CHENG, HWEI-MING  
TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
NUMBER OF SEQUENCES: 193  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MASON & ASSOCIATES, P.A.  
STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
CITY: CLEARWATER  
STATE: FLORIDA  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 3.0  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/934,915  
FILING DATE: 22-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/949,836  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LOUISE A. FOUTCH  
REGISTRATION NUMBER: 37,133  
REFERENCE/DOCKET NUMBER: 1946,6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 813-538-3800  
TELEFAX: 813-538-3820  
TELEX:

INFORMATION FOR SEQ ID NO: 163:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-934-915-163

Query Match 78.3%; Score 101; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 4.1e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCSLYGTTLE 18  
Db 3 KISEYRHVCSLYGTTLE 20

RESULT 10  
US-08-860-165-14  
Sequence 14 Application US/08860165A  
Patent No. 6004557  
GENERAL INFORMATION:  
APPLICANT: EDWARDS, Stirling John  
APPLICANT: COX, John Cooper  
APPLICANT: WEBB, Elizabeth Ann  
APPLICANT: FRAZER, Ian  
TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS  
FILE REFERENCE: 17227/130  
CURRENT APPLICATION NUMBER: US/08/860,165A  
CURRENT FILING DATE: 1997-09-22  
EARLIER APPLICATION NUMBER: PCT/AU95/00868  
EARLIER FILING DATE: 1995-12-20  
EARLIER APPLICATION NUMBER: AU PNO157  
EARLIER FILING DATE: 1994-12-20  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 14  
LENGTH: 172  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion  
US-08-860-165-14

Query Match 76.7%; Score 99; DB 3; Length 172;  
Best Local Similarity 77.3%; Pred. No. 9.3e-08;  
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KISEYRHVCSLYGTTLEQYVN 22  
Db 148 KISEYRHVCSLYGTTLSHHH 169

RESULT 11  
US-09-359-382-14  
Sequence 14 Application US/09359382  
Patent No. 6306397  
GENERAL INFORMATION:  
APPLICANT: EDWARDS, Stirling John  
APPLICANT: COX, John Cooper  
APPLICANT: WEBB, Elizabeth Ann  
APPLICANT: FRAZER, Ian  
TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS  
FILE REFERENCE: 017227/0148  
CURRENT APPLICATION NUMBER: US/09/359,382  
CURRENT FILING DATE: 1999-07-23  
EARLIER APPLICATION NUMBER: US 08/860,165  
EARLIER FILING DATE: 1997-09-22  
EARLIER APPLICATION NUMBER: PCT/AU95/00868  
EARLIER FILING DATE: 1995-12-20  
EARLIER APPLICATION NUMBER: AU PNO157/94

EARLIER FILING DATE: 1994-12-20  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 14  
LENGTH: 172  
TYPE: PRT  
ORGANISM: Human papillomavirus type 16  
US-09-359-382-14

Query Match 76.7%; Score 99; DB 4; Length 172;  
Best Local Similarity 77.3%; Pred. No. 9.3e-08;  
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KISEYRHVCSLYGTLEQQYN 22  
Db 148 KISEYRHVCSLYGTLSRHH 169

RESULT 12  
US-08-159-339A-1171  
Sequence 1171, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esteban  
TITLE OF INVENTION: HLA Binding peptides and their  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 018623-005030US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 1171:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-159-339A-1171

Query Match 51.2%; Score 66; DB 3; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00039;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 HVCYSLXGTTL 17  
Db 1 HVCYSLXGTTL 11

RESULT 13  
US-08-247-904B-10  
Sequence 10, Application US/08247904B  
Patent No. 5981699  
GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Eckstein, Jens W.  
APPLICANT: Draetta, Giulio  
TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley, Hoag & Elliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/247,904B  
FILING DATE: 23-MAY-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-247-904B-10

Query Match 45.7%; Score 59; DB 2; Length 158;  
Best Local Similarity 54.5%; Pred. No. 0.091;  
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 KISEYRHVCSLYGTLEQQYN 22  
Db 74 RIRELRHVSQVYGDLEKLTN 95

RESULT 14  
US-08-767-942A-19  
Sequence 19, Application US/08767942A  
Patent No. 6068982  
GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Chiu, M. Isabel  
APPLICANT: Berlin, Vivian  
APPLICANT: Damagnez, Veronique  
APPLICANT: Draetta, Giulio  
APPLICANT: Guillaume, Cottarel  
TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/767,942A
; FILING DATE: 17-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Mathew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-029.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-767-942A-19

```

```

Query Match      45.7%; Score 59; DB 3; Length 158;
Best Local Similarity 54.5%; Pred. No. 0.091;
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 KISEYRHVCYSLYGTTLEQQYN 22
   :| | | | | :| | | | | :|
Db 74 RIRELRHYSDSVYGDPLEKLTN 95

```

```

RESULT 15
US-08-117-083-14
; Sequence 14: Application US/08117083
; Patent No. 5719054
; GENERAL INFORMATION:
; APPLICANT: Boursnell, Michael E.
; APPLICANT: Inglis, Stephen C.
; APPLICANT: Munro, Alan J.
; TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Walter H. Dreger
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,083
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-58783
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249

```

```

; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..271
; OTHER INFORMATION: /note="Xaa refers to stop codon in
; the open reading frame."
; US-08-117-083-14

```

```

Query Match      45.7%; Score 59; DB 1; Length 271;
Best Local Similarity 54.5%; Pred. No. 0.17;
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 KISEYRHVCYSLYGTTLEQQYN 22
   :| | | | | :| | | | | :|
Db 75 RIRELRHYSDSVYGDPLEKLTN 96

```

```

Search completed: February 13, 2002, 09:59:57
Job time: 346 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 09:54:26 : Search time 39.34 seconds  
(without alignments)  
44.535 Million cell updates/sec

Title: US-09-664-225-66

Perfect score: 129

Sequence: 1 KISEYRHYCYSLVGTLEQQYKN 23

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	158	1 W6MLHS	protein E6 - human
2	91	70.5	149	1 W6ML58	E6 protein - human
3	90	69.8	148	1 A61237	E6 protein - human
4	90	69.8	148	2 S36573	E6 protein - human
5	89	69.0	149	1 W6ML33	E6 protein - human
6	85	65.9	149	1 W6ML35	E6 protein - human
7	79	61.2	149	1 W6ML31	E6 protein - human
8	68	52.7	191	1 W6MLR1	E6 protein - rhesu
9	66	51.2	150	1 W6MLC1	E6 protein - pygmy
10	66	51.2	151	1 W6ML51	E6 protein - human
11	65	50.4	150	1 W6ML11	E6 protein - human
12	65	50.4	155	1 W6ML56	E6 protein - human
13	65	50.4	155	2 A44890	E6 protein - human
14	63	48.8	150	1 W6ML6	E6 protein - human
15	63	48.8	150	2 S36544	E6 protein - human
16	62	48.1	155	1 W6ML43	E6 protein - human
17	62	48.1	148	2 S36515	E6 protein - human
18	62	48.1	150	1 W6ML44	E6 protein - human
19	61	47.3	153	1 S15621	E6 protein - human
20	61	47.3	159	1 S1561	E6 protein - human
21	61	47.3	159	2 S36497	E6 protein - human
22	59	44.7	158	1 W6ML18	E6 protein - human
23	57	44.2	153	2 S36503	E6 protein - human
24	54	41.9	158	1 W6MLPR	E6 protein - human
25	54	41.9	158	1 W6ML39	E6 protein - human
26	53	41.1	158	2 S36561	E6 protein - human
27	52	40.3	154	2 S36555	E6 protein - human
28	52	40.3	154	2 S36584	E6 protein - human
29	51	39.5	292	2 H96784	hypothetical prote

30	51	39.5	477	2 G69760	beta-glucosidase h
31	49	38.0	246	2 H36892	VI polysaccharide
32	47	36.4	680	2 A28121	major merzozolite su
33	47	36.4	1772	2 A45532	major merzozolite su
34	46.5	36.0	291	2 A36051	H+/K+-exchanging A
35	45.5	35.3	120	2 B69971	conserved hypothet
36	45.5	35.3	354	2 F36840	C15L protein - var
37	45.5	35.3	354	2 T28473	hypothetical prote
38	45.5	35.3	354	2 A72155	ELL protein - var
39	45.5	35.3	580	2 T04409	probable H+-transp
40	45.5	35.3	623	1 P4P2V9	H+-transporting At
41	45.5	35.3	623	2 S57790	H+-transporting At
42	45.5	35.3	623	2 E96818	hypothetical prote
43	45.5	35.3	1245	2 S51255	probable membrane
44	45	34.9	240	2 T35224	probable nitrate r
45	45	34.9	281	2 D71705	ribonuclease D (m

## ALIGNMENTS

```
RESULT 1
W6MLHS
protein E6 - human papillomavirus type 16
C:Species: human papillomavirus type 16
C:Date: 28-May-1986 #sequence_rev150 28-May-1986 #text_change 20-Aug-1999
C:Accession: A03682; T10427
R:Seedorf, K.; Kramer, G.; Durst, M.; Suhai, S.; Rowekamp, W.G.
Virology 145, 181-185, 1985
A:Title: Human papillomavirus type 16 DNA sequence.
A:Reference number: A22355; M01D:85246220
A:Accession: A03682
A:Molecule type: DNA
A:Residues: 1-158 <SEB>
A:Cross-references: GB:K02718; NID:q333031; PIDN:AAA46939.1; PID:q333032
R:Kennedy, I.M.; Haddow, J.K.; Clements, J.B.
J. Virol. 65, 2093-2097, 1991
A:Title: A negative element in the human papillomavirus type 16 genome acts at the 1
A:Reference number: Z17014; M01D:91162763
A:Accession: T10427
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-158 <KEN>
A:Cross-references: EMBL:K02718; NID:q333031; PIDN:AAA46939.1; PID:q333032
C:Genetics:
A:Gene: E6
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; zinc finger
F:37-73/Region: zinc finger CCCC motif
F:110-146/Region: zinc finger CCCC motif

Query Match 100.0%; Score 129; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 5,6e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHYCYSLVGTLEQQYKN 23
DB 79 KISEYRHYCYSLVGTLEQQYKN 101

RESULT 2
W6ML58
E6 protein - human papillomavirus type 58
C:Species: human papillomavirus type 58
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_rev150 31-Mar-1992 #text_change 21-Jul-2000
C:Accession: E36779
R:Kirii, Y.; Imamoto, S.; Matsukura, T.
Virology 185, 424-427, 1991
A:Title: Human papillomavirus type 58 DNA sequence.
A:Reference number: A36779; M01D:92024102
A:Accession: E36779
```

A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-149 <KIR>  
A:Cross-references: GB:D90400; NID:g222386; PIDN:BAA31845.1; PID:g3337098  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999  
C:Accession: A03683  
R:Coile, S.T.; Streck, R.E.  
J. Virol. 58, 991-995, 1986  
A:Title: Genome organization and nucleotide sequence of human papillomavirus type 33,  
A:Reference number: A93020; MUID:86200464  
A:Accession: A03683  
A:Molecule type: DNA  
A:Residues: 1-149 <COL>  
A:Cross-references: GB:M12732; NID:g333049; PIDN:AAA46958.1; PID:g463177  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; zinc finger  
F:30-66/Region: zinc finger CCCC motif  
F:103-139/Region: zinc finger CCCC motif

Query Match 70.5%; Score 91; DB 1; Length 149;  
Best Local Similarity 78.3%; Pred. No. 2.3e-06;  
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KISEYRHYCYSLYGTLEQOYK 23  
||||| ||||| |||||  
DB 72 KISEYRHYCYSLYGTLEQOYK 94

RESULT 3  
A61237  
E6 protein - human papillomavirus type 52  
C:Species: human papillomavirus type 52  
C:Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 03-May-1996  
C:Accession: A61237  
R:Takami, Y.; Kondo, G.; Saito, J.; Noda, K.; Sudiro, T.M.; Sjahrurachman, A.; Warsa, U  
Int. J. Cancer 48, 516-522, 1991  
A:Title: Cloning and characterization of human papillomavirus type 52 from cervical carc  
A:Reference number: A61237; MUID:91258022  
A:Accession: A61237  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-148 <TRK>  
C:Superfamily: papillomavirus E6 protein

Query Match 69.8%; Score 90; DB 2; Length 148;  
Best Local Similarity 73.9%; Pred. No. 3.3e-06;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KISEYRHYCYSLYGTLEQOYK 23  
||||| ||||| ||||| |||||  
DB 72 KISEYRHYCYSLYGTLEQOYK 94

RESULT 4  
S36573  
E6 protein - human papillomavirus type 52  
C:Species: human papillomavirus type 52  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S36573  
R:Deilus, H.; Hofmann, B.  
submitted to the EMBL data library, August 1993  
A:Description: Primer-directed sequencing of human papillomavirus types.  
A:Reference number: S36469  
A:Accession: S36573  
A:Molecule type: DNA  
A:Residues: 1-148 <DEL>  
A:Cross-references: EMBL:X74481; NID:g397038; PIDN:CAA52585.1; PID:g397039  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 69.8%; Score 90; DB 2; Length 148;  
Best Local Similarity 73.9%; Pred. No. 3.3e-06;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KISEYRHYCYSLYGTLEQOYK 23  
||||| ||||| ||||| |||||  
DB 72 KISEYRHYCYSLYGTLEQOYK 94

RESULT 5

M6WL33  
E6 protein - human papillomavirus type 33  
C:Species: human papillomavirus type 33  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999  
C:Accession: A03683  
R:Coile, S.T.; Streck, R.E.  
J. Virol. 58, 991-995, 1986  
A:Title: Genome organization and nucleotide sequence of human papillomavirus type 33,  
A:Reference number: A93020; MUID:86200464  
A:Accession: A03683  
A:Molecule type: DNA  
A:Residues: 1-149 <COL>  
A:Cross-references: GB:M12732; NID:g333049; PIDN:AAA46958.1; PID:g463177  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; zinc finger  
F:30-66/Region: zinc finger CCCC motif  
F:103-139/Region: zinc finger CCCC motif

Query Match 69.0%; Score 89; DB 1; Length 149;  
Best Local Similarity 73.9%; Pred. No. 4.7e-06;  
Matches 17; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 KISEYRHYCYSLYGTLEQOYK 23  
||||| ||||| ||||| |||||  
DB 72 KISEYRHYCYSLYGTLEQOYK 94

RESULT 6  
M6WL35  
E6 protein - human papillomavirus type 35  
C:Species: human papillomavirus type 35  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Mar-2001  
C:Accession: E40824; S36521  
R:Marich, J.E.; Pontier, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.  
Virology 186, 770-776, 1992  
A:Title: The phylogenetic relationship and complete nucleotide sequence of human pap  
A:Reference number: A40824; MUID:92124753  
A:Accession: E40824  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-149 <MAR>  
A:Cross-references: GB:M74117; NID:g333050; PIDN:AAA46966.1; PID:g333051  
R:Deilus, H.; Hofmann, B.  
submitted to the EMBL data library, August 1993  
A:Description: Primer-directed sequencing of human papillomavirus types.  
A:Reference number: S36469  
A:Accession: S36521  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-149 <DEL>  
A:Cross-references: EMBL:X74477; NID:g396997; PIDN:CAA52561.1; PID:g396998  
A:Experimental source: strain 35H  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; zinc finger  
F:30-66/Region: zinc finger CCCC motif  
F:103-139/Region: zinc finger CCCC motif

Query Match 65.9%; Score 85; DB 1; Length 149;  
Best Local Similarity 73.9%; Pred. No. 1.8e-05;  
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KISEYRHYCYSLYGTLEQOYK 23  
||||| ||||| ||||| |||||  
DB 72 KISEYRHYCYSLYGTLEQOYK 94

RESULT 7  
M6WL31  
E6 protein - human papillomavirus type 31  
C:Species: human papillomavirus type 31



## RESULT 12

W6WLL6

E6 protein - human papillomavirus type 56

C:Species: human papillomavirus type 56

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999

C:Accession: A33377; S36579

R:Loerincz, A.T.; Quinn, A.P.; Goldsbrough, M.D.; McAllister, P.; Temple, G.F.

J. Gen. Virol. 70, 3099-3104, 1989

A:Title: Human papillomavirus type 56: a new virus detected in cervical cancers.

A:Reference number: A33377; MUID:90063558

A:Accession: A33377

A:Molecule type: DNA

A:Residues: 1-155 &lt;LOE&gt;

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36579

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 &lt;DEL&gt;

C:Cross-references: EMBL:X74483; NID:9397053; PID:CAA52596.1; PID:9397054

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; transforming protein; zinc finger

F:33-69/Region: zinc finger CCCC motif

F:106-142/Region: zinc finger CCCC motif

Query Match

Best Local Similarity 50.4%; Score 65; DB 1; Length 155;

Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLYGTLE 18

Db 75 KVRKRYRYDYSVYGTLE 92

## RESULT 13

A44890

E6 protein - human papillomavirus type 66

C:Species: human papillomavirus type 66

C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 22-May-1998

C:Accession: A44890

R:Twined, A.R.; Beaudenon, S.; Favre, M.; Orth, G.

J. Clin. Microbiol. 29, 2656-2660, 1991

A:Title: Characterization of human papillomavirus type 66 from an invasive carcinoma of

A:Reference number: A44890; MUID:92129556

A:Accession: A44890

A:Molecule type: DNA

A:Residues: 1-155 &lt;TAM&gt;

A:Note: sequence extracted from NCBI backbone (NCBI:78637, NCBI:78638)

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; transforming protein; zinc finger

F:33-69/Region: zinc finger CCCC motif

F:106-142/Region: zinc finger CCCC motif

Query Match

Best Local Similarity 50.4%; Score 65; DB 2; Length 155;

Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLYGTLE 18

Db 75 KVRKRYRYDYSVYGTLE 92

## RESULT 14

W6WLL6

E6 protein - human papillomavirus type 6b

C:Species: human papillomavirus type 6b

C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999

C:Accession: E20558

R:Schwarz, E.; Durst, M.; Demankowski, C.; Lattermann, O.; Zech, R.; Wolfesperger, E.;

EMBO J. 2, 2341-2348, 1983

A:Title: DNA sequence and genome organization of genital human papillomavirus type 6b

A:Reference number: A90975; MUID:84133949

A:Accession: E20558

A:Molecule type: DNA

A:Residues: 1-150 &lt;SCH&gt;

C:Cross-references: GB:X00203; NID:960955; PID:CAA25018.1; PID:960956

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; zinc finger

F:31-67/Region: zinc finger CCCC motif

F:104-140/Region: zinc finger CCCC motif

Query Match

Best Local Similarity 48.8%; Score 63; DB 1; Length 150;

Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLYGTLE 20

Db 73 KINCYRHFDYAGYATVVEE 92

## RESULT 15

S36544

E6 protein - human papillomavirus type 26

C:Species: human papillomavirus type 26

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36544

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36544

A:Molecule type: DNA

A:Residues: 1-150 &lt;DEL&gt;

A:Cross-references: EMBL:X74472; NID:9396956; PID:CAA52530.1; PID:9396957

C:Superfamily: papillomavirus E6 protein

C:Keywords: early protein; zinc finger

Query Match

Best Local Similarity 48.8%; Score 63; DB 2; Length 150;

Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLYGTLE 18

Db 72 KITERYRTCSYVGTLE 89

Search completed: February 13, 2002, 10:00:43  
job time: 377 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 09:58:11 : Search time 21.6 seconds

(without alignments)  
39.041 Million cell updates/sec

Title: US-09-664-225-66

Perfect score: 129  
Sequence: 1 KISEYRHVCYSLVCTLEQYXNK 23

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SWISSPROT\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	129	100.0	158	VE6_HPV16	P03126 human papill
2	91	70.5	149	VE6_HPV58	P26585 human papill
3	90	69.8	148	VE6_HPV52	P36814 human papill
4	89	69.0	149	VE6_HPV33	P06437 human papill
5	85	65.9	149	VE6_HPV35	P27228 human papill
6	79	61.2	149	VE6_HPV31	P17386 human papill
7	68	52.7	191	VE6_HPV1	P22139 thesus papill
8	66	51.2	150	VE6_PCPV1	Q02270 pygmy chimp
9	66	51.2	150	VE6_HPV51	P26554 human papill
10	65	50.4	150	VE6_HPV11	P04019 human papill
11	65	50.4	155	VE6_HPV56	P24836 human papill
12	65	50.4	155	VE6_HPV66	O80955 human papill
13	63	48.8	150	VE6_HPV26	P36807 human papill
14	63	48.8	150	VE6_HPV6A	O84291 human papill
15	63	48.8	150	VE6_HPV6B	P06462 human papill
16	63	48.8	155	VE6_HPV43	P19762 human papill
17	62	48.1	148	VE6_HPV44	P36811 human papill
18	62	48.1	150	VE6_HPV34	P19710 human papill
19	61	47.3	153	VE6_HPV2A	P25464 human papill
20	61	47.3	153	VE6_HPV57	P22138 human papill
21	61	47.3	159	VE6_HPV27	P36808 human papill
22	60	46.5	158	VE6_HPV70	P50804 human papill
23	59	45.7	158	VE6_HPV18	P06463 human papill
24	57	44.2	153	VE6_HPV30	P36809 human papill
25	56	43.4	146	VE6_HPV61	O80948 human papill
26	54	41.9	158	VE6_HPV39	P24835 human papill
27	54	41.9	158	VE6_HPV68	P54667 human papill
28	54	41.9	158	VE6_HPV6E	P27962 human papill
29	53	41.1	158	VE6_HPV45	P21735 human papill
30	52	40.3	154	VE6_HPV07	P36800 human papill
31	52	40.3	154	VE6_HPV40	P36812 human papill
32	51	39.5	477	BGL2_SACTSU	P42403 bacillus su
33	49	38.0	246	VE6C_SALTI	P43110 salmonella

## ALIGNMENTS

34	47	36.4	1772	1	MSPI_PLAYO	P13828 plasmodium
35	46.5	36.0	291	1	ATRB_RABIT	P18597 oryctolagus
36	46	35.7	1138	1	C7AB_BACUA	O45707 bacillus th
37	46	35.7	1138	1	C7AB_BACUK	O45708 bacillus th
38	46	35.7	1167	1	C1JA_BACTU	O45738 bacillus th
39	45.5	35.3	120	1	YRAJ_BACST	O07934 bacillus su
40	45.5	35.3	354	1	VE11_VARY	P33870 variola vir
41	45.5	35.3	580	1	VATP_HORVU	O40002 bordum vul
42	45.5	35.3	623	1	VATP_ARATH	O23654 arabidopsis
43	45.5	35.3	623	1	VATP_BRANA	O39291 brassica na
44	45.5	35.3	623	1	VATP_DAUCA	P09469 daucus caro
45	45	34.9	337	1	VE11_VACCP	P29888 vaccinia vl

RESULT 1  
ID VE6\_HPV16 STANDARD; PRT; 158 AA.  
AC P03126;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE E6 PROTEIN.  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85246220; PubMed=2990099;  
RA Seedorf K., Krammer G., Burst W., Suhai S., Roweckamp W.G.;  
RT "Human papillomavirus type 16 DNA sequence.";  
RL Virology 145:181-185(1985).  
RN [2]  
RP SEQUENCE OF 31-50 FROM N.A.  
RX MEDLINE=90218027; PubMed=2157796;  
RA Schneider-Mannoury S., Pehau-Arnaut G., Breilburd F., Orth G.;  
RT "Expression of the human papillomavirus type 16 genome in SK-V cells, a line derived from a vulvar intraepithelial neoplasia.";  
RL J. Gen. Virol. 71:809-817(1990).  
CC -!- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.  
CC -!- FUNCTION: EXHIBITS A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE STRANDED DNA (IN VITRO).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.  
CC -!- MISCELLANEOUS: HPV16, IN COMPARISON TO HPV TYPES 6 AND 11, IS MORE OFTEN ASSOCIATED WITH MALIGNANT GENITAL CANCERS IN HUMANS.  
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: K02718; AAA46939.1; -;  
DR EMBL: D00735; BAA00632.1; -;  
DR PIR: A03682; MWLMS.  
DR InterPro: IPR00134; E6.  
DR Pfam: PF00518; E6; 1.  
KW Early protein; DNA-binding; Nuclear protein; zinc-finger;  
KW Transforming protein.  
FT ZN\_FING 37  
FT ZN\_FING 110 146 POTENTIAL.  
SQ SEQUENCE 158 AA; 19187 MW; 01FE5ADCDFB37EB CRC64;  
Query Match 100.0%; Score 129; DB 1; Length 156;  
Best local Similarity 100.0%; Pred. No. 4, 2e-12;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      1 KISEYRHVCYSLYGTTLEQOYRK 23
      |||||||
Db      79 KISEYRHVCYSLYGTTLEQOYRK 101

RESULT  2
VE6_HPVS2  STANDARD;  PRT;  149 AA.
ID  VE6_HPVS2
AC  P36814;
DT  01-AUG-1992 (Rel. 23, Created)
DT  01-AUG-1992 (Rel. 23, Last sequence update)
DT  01-JUN-1994 (Rel. 29, Last annotation update)
DE  E6 PROTEIN.
GN  E6.
OS  Human papillomavirus type 58.
OC  Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;
CC  Papillomavirus.
CC  NCBI_TaxID=10598;
CC  [1]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=92024102; PubMed=1656594;
RA  Kiril Y., Iwamoto S., Matsukura T.;
RT  *Human papillomavirus type 58 DNA sequence.*;
RL  Virology 185:424-427(1991).
CC  -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC  STRANDED DNA (IN VITRO).
CC  -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; D90400; BA31845.1; -.
DR  PIR; E36779; W6WL58.
DR  InterPro; IPR001334; E6.
DR  Pfam; PF00518; E6; 1.
KM  Early protein: DNA-binding; Nuclear protein; Zinc-finger.
FT  ZN_FING 30 66 POTENTIAL.
FT  ZN_FING 103 139 POTENTIAL.
SQ  SEQUENCE 149 AA: 17794 MW: 79B3DC95831B158 CRC64;

Query Match 70.5%; Score 91; DB 1; Length 149;
Best Local Similarity 78.3%; Pred. No. 1.5e-06;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY      1 KISEYRHVCYSLYGTTLEQOYRK 23
      |||||||
Db      72 KISEYRHVCYSLYGTTLEQOYRK 94

RESULT  3
VE6_HPVS2  STANDARD;  PRT;  148 AA.
ID  VE6_HPVS2
AC  P36814;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  E6 PROTEIN.
GN  E6.
OS  Human papillomavirus type 52.
OC  Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;
CC  Papillomavirus.
CC  NCBI_TaxID=10618;
CC  [1]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=94265501; PubMed=8205838;
RA  Delius H., Hofmann B.;

```

```

RT  "Primer-directed sequencing of human papillomavirus types.";
RL  Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC  -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC  STRANDED DNA (IN VITRO).
CC  -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; X74481; CAA52585.1; -.
DR  PIR; S36573; S36573.
DR  InterPro; IPR001334; E6.
DR  Pfam; PF00518; E6; 1.
KM  Early protein: DNA-binding; Nuclear protein; Zinc-finger.
FT  ZN_FING 30 66 POTENTIAL.
FT  ZN_FING 103 139 POTENTIAL.
SQ  SEQUENCE 148 AA: 17698 MW: C5E9DCE341AB0DE CRC64;

Query Match 69.8%; Score 90; DB 1; Length 148;
Best Local Similarity 73.9%; Pred. No. 2e-06;
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY      1 KISEYRHVCYSLYGTTLEQOYRK 23
      |||||||
Db      72 KISEYRHVCYSLYGTTLEQOYRK 94

RESULT  4
VE6_HPVS3  STANDARD;  PRT;  149 AA.
ID  VE6_HPVS3
AC  P06427;
DT  01-JAN-1988 (Rel. 06, Created)
DT  01-JAN-1988 (Rel. 06, Last sequence update)
DT  01-JUN-1994 (Rel. 29, Last annotation update)
DE  E6 PROTEIN.
GN  E6.
OS  Human papillomavirus type 33.
OC  Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;
CC  Papillomavirus.
CC  NCBI_TaxID=10586;
CC  [1]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=86200464; PubMed=3009902;
RA  Cole S.T., Strecek R.E.;
RT  *Genome organization and nucleotide sequence of human papillomavirus
RT  type 33, which is associated with cervical cancer.*;
RL  J. Virol. 58:991-995(1986).
CC  -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE ONCOGENIC POTENTIAL
CC  OF THIS VIRUS (CERVICAL NEOPLASTIA-ASSOCIATED VIRUS).
CC  -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC  STRANDED DNA (IN VITRO).
CC  -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; M12732; AAA46958.1; -.
DR  PIR; A03683; W6WL33.
DR  InterPro; IPR001334; E6.
DR  Pfam; PF00518; E6; 1.
KM  Early protein: DNA-binding; Nuclear protein; Zinc-finger;
RN  Transforming protein.

```





```

OS Human papillomavirus type 11.
OC Viruses: dsDNA viruses, no RNA stage: Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=65181601; Pubmed=3008427;
RA Datmann K., Schwarz E., Gissmann L., Zur Hausen H.;
RT "The nucleotide sequence and genome organization of human papilloma
virus type 11."
RL Virology 151:124-130(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA File K.H., Fan L., Fritsch M.H., Bryan J., Brown D.R.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON-SPECIFIC AFFINITY FOR DOUBLE
STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/sib-sib.ch).
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M14119; AAA46927.1; -.
DR EMBL: L36108; AAA21703.1; -.
DR PIR: A03684; W6W111.
DR InterPro: IPR001334; EG.
DR Pfam: PF00518; Eg_1.
DR K: Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT FT_FING 31 67
FT ZN_FING 104 140 POTENTIAL.
FT FT_FING 150 AA; 17406 MW; A685C7DCA090CCA CRC64;
SO SEQUENCE

*
Query Match 50.4%; Score 65; DB 1; Length 150;
Best Local Similarity 43.5%; Pred. No. 0.0094;
Matches 10; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 KISEYRHVYCYSLYGTTLEQOYVK 23
DB 73 KINQYRHFNVAAYAPVEETNE 95
II:||||:|:|:|:|:|
73 KINQYRHFNVAAYAPVEETNE 95

RESULT 11
ID_V66_HPV56 STANDARD: PRT; 155 AA.
AC P24836;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 56.
OC Viruses: dsDNA viruses, no RNA stage: Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10596;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90063558; Pubmed=2555440;
RA Loerincz A.T., Quinn A.P., Goldsborough M.D., McAllister P.,
RT "Human papillomavirus type 56: a new virus detected in cervical
cancers."
RL J. Gen. Virol. 70:3099-3104(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; Pubmed=8205838;
RA Delius H., Hofmann B.;

```

```

RT "Primer-directed sequencing of human papillomavirus types."
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/or\_send\_an\_email\_to\_license@isb-sib.ch).
CC -----
DR EMBL; X74483; CAA52596.1; -.
DR PIR; A33377; W6MNL56.
DR PIR; S36579; S36579.
DR HSSP; P00277; 1EGR.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;
KW Transforming protein.
FT ZN_FING 33 69 POTENTIAL.
FT ZN_FING 106 142 POTENTIAL.
FT SEQUENCE 155 AA; 18524 MW; E8732949398B4C4B CRC64;

Query Match 50.4%; Score 65; DB 1; Length 155;
Best Local Similarity 61.1%; Pred. NO. 0.0097;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLSYGTLE 18
   1: :||: ||||| |||
Db 75 KYRKYRYDYSGYATLE 92

RESULT 12
VE6_HPv66
ID VE6_HPv66 STANDARD; PRT; 155 AA.
AC Q80955;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 66.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=37119;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/or\_send\_an\_email\_to\_license@isb-sib.ch).
CC -----
DR EMBL; U31794; AAA79499.1; -.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT ZN_FING 33 69 BY SIMILARITY.
FT ZN_FING 106 142 BY SIMILARITY.
FT SEQUENCE 155 AA; 18609 MW; 905841B1A35E8355 CRC64;

```



KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
FT ZN\_FING 31 67 POTENTIAL.  
FT ZN\_FING 104 140 POTENTIAL.  
SQ SEQUENCE 150 AA; 17299 MW; F7CSEF80356A4DF9 CRC64;

Query Match 48.8%; Score 63; DB 1; Length 150;  
Best Local Similarity 50.0%; Pred. No. 0.018;  
Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 KISEYRHCYSLYGTTLEQQ 20  
||:||||:|:| ||:|:  
DB 73 KINQYRHRFDYAGYATVVEE 92

Search completed: February 13, 2002, 10:02:20  
Job time: 249 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 09:57:41 ; Search time 62.52 Seconds  
(without alignments)  
53.811 Million cell updates/sec

Title: US-09-664-225-66  
Perfect score: 129  
Sequence: 1 KISEYRHYCYSKGTLEEQYKN 23

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 473505 seqs, 14627329 residues  
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP-REMBL\_17:\*  
2: sp-archaea:\*  
3: sp-bacteria:\*  
4: sp-fungi:\*  
5: sp-human:\*  
6: sp-invertebrate:\*  
7: sp-mammal:\*  
8: sp-mhc:\*  
9: sp-organellar:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	90	12	080883 human papill
2	129	100.0	90	12	080884 human papill
3	129	100.0	151	12	089640 human papill
4	129	100.0	151	12	012336 human papill
5	129	100.0	151	12	09WMP4 human papill
6	129	100.0	151	12	09W931 human papill
7	129	100.0	151	12	089755 human papill
8	129	100.0	158	12	09QDH9 human papill
9	129	100.0	158	12	09QDH7 human papill
10	129	100.0	158	12	09QDH5 human papill
11	129	100.0	158	12	09QDH3 human papill
12	126	97.7	84	12	080882 human papill
13	126	97.7	151	12	089648 human papill
14	126	97.7	151	12	080963 human papill
15	126	97.7	151	12	012335 human papill
16	126	97.7	151	12	09WMP5 human papill
17	126	97.7	151	12	089852 human papill
18	126	97.7	158	12	09WH13 human papill
19	123	95.3	90	12	080885 human papill

20	123	95.3	151	12	09WMP2 human papill
21	123	95.3	151	12	09W8C3 human papill
22	123	95.3	151	12	089708 human papill
23	123	95.3	151	12	080966 human papill
24	120	93.0	81	12	080886 human papill
25	120	93.0	91	12	080887 human papill
26	120	93.0	151	12	089887 human papill
27	120	93.0	151	12	09WMP3 human papill
28	97	75.2	149	12	091ZG6 human papill
29	92	71.3	149	12	091ZG5 human papill
30	87	67.4	149	12	084298 human papill
31	74	57.4	149	12	090723 human papill
32	72	55.8	148	12	082005 human papill
33	69	53.5	152	12	037385 common chim
34	63	48.8	151	12	091RS0 human papill
35	63	48.8	151	12	0994A0 human papill
36	60	46.5	151	12	091H51 human papill
37	59	45.7	148	12	081997 human papill
38	59	45.7	158	12	09QNP8 human papill
39	57	44.2	147	12	09WNN0 human papill
40	56	43.4	157	12	09WHG0 human papill
41	55	42.6	148	12	099FX3 human papill
42	55	42.6	150	12	098004 human papill
43	54	41.9	162	4	0901Z9 homo sapien
44	54	41.9	717	11	09WTX8 mus musculu
45	53	41.1	158	4	09Y4Y4 homo sapien

## ALIGNMENTS

RESULT 1  
ID 080883 PRELIMINARY; PRT: 90 AA.  
AC 080883;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE E6 PROTEIN (FRAGMENT).  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16;  
RA Haegert D.G., Galutira D.F., Younghusband B.H.;  
RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.  
DR EMBL: U14512; AAB60566.1; .  
DR InterPro: IPRO01334; E6.  
DR Pfam: PF00518; E6; 1.  
FT NON\_TER 1 1  
FT 90 90  
SQ SEQUENCE 90 AA; 10904 MW; 5D3ADF843AD6060B CRC64;  
  
Query Match 100.0%; Score 129; DB 12; Length 90;  
Best Local Similarity 100.0%; Pred. No. 8.2e-13;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 KISEYRHYCYSKGTLEEQYKN 23  
DB 31 KISEYRHYCYSKGTLEEQYKN 53  
  
RESULT 2  
ID 080884 PRELIMINARY; PRT: 90 AA.  
AC 080884;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE E6 PROTEIN (FRAGMENT).  
 GN E6.  
 OS Human Papillomavirus type 16.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10581;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HPV16;  
 RA Haegert D.G., Galutira D.F., Youngusband B.H.;  
 RT "Sequence variation in the E6 gene of human papillomavirus type 16."  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U14513; AAB60567.2; -;  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 90 90  
 SO SEQUENCE 90 AA; 11021 MW; 47F42BBERACCC01 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 8, 2e-13;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLVGTTLTLEQYKN 23  
 ||||||||||||||||||||  
 DB 31 KISEYRHVCYSLVGTTLTLEQYKN 53

RESULT 3  
 ID 089640 PRELIMINARY; PRT; 151 AA.  
 AC 089640;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 31-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE EARLY TRANSDOMING PROTEIN E6.  
 GN E6.  
 OS Human Papillomavirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10566;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OR 7574;  
 RX MEDLINE=96079021; PubMed=7494284;  
 RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,  
 RA Jenison S.A.;  
 RT "Human papillomavirus type 16 variant lineages in United States  
 RT populations characterized by nucleotide sequence analysis of the E6,  
 RT L2, and L1 coding segments."  
 RL J. Virol. 69:7743-7753(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OR 7574;  
 RA Farmer A.D.;  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U34129; AAA91676.1; -;  
 DR EMBL: U34115; AAA91662.1; -;  
 DR EMBL: U34120; AAA91667.1; -;  
 DR EMBL: U34124; AAA91671.1; -;  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 SO SEQUENCE 151 AA; 18348 MW; FE3F2A2FCFOA6CB2 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1, 4e-12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLVGTTLTLEQYKN 23  
 ||||||||||||||||||||  
 DB 72 KISEYRHVCYSLVGTTLTLEQYKN 94

RESULT 4  
 ID 012336 PRELIMINARY; PRT; 151 AA.  
 AC 012336;  
 DT 01-JUL-1997 (TREMblrel. 04, Created)  
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE E6 PROTEIN.  
 OS Human Papillomavirus type 16.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10581;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RC-7;  
 RX MEDLINE=97437474; PubMed=9292007;  
 RA Tornesello M.L., Buongiorno F.M., Meglio A., Buongiorno L.,  
 RA Beth-Girardo E., Girardo G.;  
 RT "Sequence variations and viral genomic state of human papillomavirus  
 RT type 16 in penile carcinomas from Ugandan patients."  
 RL J. Gen. Virol. 78:0-0(0).  
 DR EMBL: AF003016; AAB70733.1; -;  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 SO SEQUENCE 151 AA; 18206 MW; 51C12A8B149C6D8A CRC64;

Query Match 100.0%; Score 129; DB 12; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1, 4e-12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLVGTTLTLEQYKN 23  
 ||||||||||||||||||||  
 DB 72 KISEYRHVCYSLVGTTLTLEQYKN 94

RESULT 5  
 ID 09WMP4 PRELIMINARY; PRT; 151 AA.  
 AC 09WMP4;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE E6 PROTEIN.  
 GN E6.  
 OS Human Papillomavirus type 16.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10581;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BC410T;  
 RA van Duijn M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,  
 RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;  
 RT "Analysis of Human Papillomavirus type 16 E6 variants in relation to  
 RT p53 codon 72 polymorphism genotype in cervical carcinogenesis."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ388060; CAB45112.1; -;  
 DR InterPro: IPR001334; E6.  
 DR Pfam: PF00518; E6; 1.  
 SO SEQUENCE 151 AA; 18387 MW; E2244784BBA6C02 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1, 4e-12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLVGTTLTLEQYKN 23  
 ||||||||||||||||||||  
 DB 72 KISEYRHVCYSLVGTTLTLEQYKN 94

```

RESULT 6
O9W931 PRELIMINARY; PRT; 151 AA.
AC O9W931;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EA2461.2;
RA Van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;
RT *Analysis of Human papillomavirus type 16 E6 variants in relation to
RT p53 codon 72 polymorphism genotype in cervical carcinogenesis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ388058; CAB45110.1; -.
DR EMBL; AJ388058; CAB45108.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18306 MW; 6FB3D9E0F24A5300 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTLEOQYNK 23
Db 72 KISEYRHVCYSLYGTLEOQYNK 94
|||||
O989755 PRELIMINARY; PRT; 151 AA.
AC O989755;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE EARLY TRANSFORMING PROTEIN E6.
GN E6.
OS Human papillomavirus, and
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Human papillomavirus; STRAIN-OR 6170;
RX MEDLINE-96079021; PubMed-7494284;
RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
RA Jensen S.A.;
RT *Human papillomavirus type 16 variant lineages in United States
RT populations characterized by nucleotide sequence analysis of the E6,
RT L2, and L1 coding segments.";
RL J. Virol. 69:7743-7753(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Human papillomavirus; STRAIN-OR 6170;
RA Farmer A.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-Human papillomavirus type 16; STRAIN-AS, EA7768T, AND EA7496T;
RA Van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;
RT *Analysis of Human papillomavirus type 16 E6 variants in relation to

```

```

RT p53 codon 72 polymorphism genotype in cervical carcinogenesis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34126; AAA91673.1; -.
DR EMBL; U34121; AAA91658.1; -.
DR EMBL; U34121; AAA91668.1; -.
DR EMBL; U34123; AAA91670.1; -.
DR EMBL; AJ388066; CAB45124.1; -.
DR EMBL; AJ388056; CAB45104.1; -.
DR EMBL; AJ388061; CAB45114.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18334 MW; FFBF2A2FCBAC02 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTLEOQYNK 23
Db 72 KISEYRHVCYSLYGTLEOQYNK 94
|||||
O9Q0DH9 PRELIMINARY; PRT; 158 AA.
AC O9Q0DH9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KE6-1;
RA Lee H.P., Song Y.S., Kim J.W., Roh J.W., Park N.H., Kang S.B.;
RT "The most commonly found Human Papillomavirus type 16 E6 and E7
RT sequence variation tend to occur together.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187866; AAF13392.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 158 AA; 19187 MW; 15BFB5ADC6B36F0 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.5e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISEYRHVCYSLYGTLEOQYNK 23
Db 79 KISEYRHVCYSLYGTLEOQYNK 101
|||||
O9Q0DH7 PRELIMINARY; PRT; 158 AA.
AC O9Q0DH7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-KE6-2;
RA Lee H.P., Song Y.S., Kim J.W., Roh J.W., Park N.H., Kang S.B.;
RT "The most commonly found Human Papillomavirus type 16 E6 and E7
sequence variation tend to occur together.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF187867; AAF13394.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 158 AA; 19215 MW; FCSBFB06576864B CRC64;

Query Match
Best Local Similarity 100.0%; Score 129; DB 12; Length 158;
Pred. No. 1.5e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLYGTTLEQOYNK 23
Db 79 KISEYRHVCYSLYGTTLEQOYNK 101

RESULT 10
Q90DH5 PRELIMINARY; PRT; 158 AA.
AC Q90DH5;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
CC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KE6-3;
RA Lee H.P., Song Y.S., Kim J.W., Roh J.W., Park N.H., Kang S.B.;
RT "The most commonly found Human Papillomavirus type 16 E6 and E7
sequence variation tend to occur together.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF187868; AAF13396.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
DR NON_TER 84
SQ SEQUENCE 158 AA; 19195 MW; 00564E1A8994CB0B CRC64;

Query Match
Best Local Similarity 100.0%; Score 129; DB 12; Length 158;
Pred. No. 1.5e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLYGTTLEQOYNK 23
Db 79 KISEYRHVCYSLYGTTLEQOYNK 101

RESULT 11
Q90DH3 PRELIMINARY; PRT; 158 AA.
AC Q90DH3;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE E6 PROTEIN.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
CC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KE6-4;
RA Lee H.P., Song Y.S., Kim J.W., Roh J.W., Park N.H., Kang S.B.;
RT "The most commonly found Human Papillomavirus type 16 E6 and E7

```

```

RT sequence variation tend to occur together.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF187869; AAF13398.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
DR NON_TER 84
SQ SEQUENCE 158 AA; 19210 MW; 004EF5ADD6FABEB3B CRC64;

Query Match
Best Local Similarity 100.0%; Score 129; DB 12; Length 158;
Pred. No. 1.5e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLYGTTLEQOYNK 23
Db 79 KISEYRHVCYSLYGTTLEQOYNK 101

RESULT 12
Q80882 PRELIMINARY; PRT; 84 AA.
AC Q80882;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE E6 PROTEIN (FRAGMENT).
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
CC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HPV16;
RA Haegert D.G., Galutira D.F., Younghusband B.H.;
RT "Sequence variation in the E6 gene of human papillomavirus type 16.";
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U14511; AAB00565.2; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
DR NON_TER 1
FT VARIANT 39 39 V -> L.
FT NON_TER 84
SQ SEQUENCE 84 AA; 10177 MW; 5AB6B896468E1CAA CRC64;

Query Match
Best Local Similarity 97.7%; Score 126; DB 12; Length 84;
Pred. No. 2.2e-12;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISEYRHVCYSLYGTTLEQOYNK 23
Db 28 KISEYRHVCYSLYGTTLEQOYNK 50

RESULT 13
Q89648 PRELIMINARY; PRT; 151 AA.
AC Q89648;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE EARLY TRANSFORMING PROTEIN E6.
GN E6.
OS Human papillomavirus, and
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
CC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OR 9237;
RX MEDLINE=06079021; Pubmed=7494284;
RA Yamada T., Wheeler C.W., Halpern A.L., Stewart A.C., Hildesheim A.,

```



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 10:02:20 ; Search time 21.6 seconds

(without alignments)  
32.251 Million cell updates/sec

Title: US-09-664-225-154

Sequence: 1 KATLQDIVLLEPQNEIPV 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	105	1	VE7_HPV18
2	71.5	74.5	106	1	VE7_HPV45
3	55	57.3	105	1	VE7_HPV56
4	54	56.2	101	1	VE7_HPV51
5	52.5	54.7	109	1	VE7_HPV39
6	51.5	53.6	109	1	VE7_HPV70
7	50.5	52.6	110	1	VE7_HPV68
8	50	52.1	105	1	VE7_HPV66
9	48.5	50.5	110	1	VE7_HPV66
10	47	49.0	803	1	SMI6_YEAST
11	46	47.9	249	1	RS3_SCHPO
12	45	46.9	113	1	VE7_HRPV1
13	45	45.8	901	1	VE7_HRPV1
14	45	45.8	95	1	VE7_HRPV54
15	44	45.8	96	1	VE7_HRPV24
16	44	45.8	97	1	VE7_HRPV34
17	44	45.8	99	1	VE7_HRPV35
18	44	45.8	954	1	DRP2_HUMAN
19	44	45.8	1616	1	VE7_HRPV52
20	43	44.8	99	1	VE7_HRPV52
21	43	44.8	1613	1	VIR2_CAEEL
22	42	43.8	97	1	VE7_HRPV44
23	42	43.8	98	1	VE7_HRPV11
24	42	43.8	98	1	VE7_HRPV6A
25	42	43.8	98	1	VE7_HRPV6B
26	42	43.8	255	1	GLIC_MOUSE
27	42	43.8	490	1	YIHF_ECOLI
28	42	43.8	634	1	KPC3_DROME
29	42	43.8	5032	1	RYNR_HUMAN
30	41.5	43.2	170	1	LXNS_ECOLI
31	41	42.7	81	1	HFO_CLOAB
32	41	42.7	98	1	VE7_HRPV31
33	41	42.7	126	1	US05_HCMVA

## ALIGNMENTS

RESULT	ID	STANDARD	PRT	105 AA.	
34	41	42.7	187	1	Y503_TREPA
35	41	42.7	829	1	CAD3_HUMAN
36	41	42.7	880	1	RPAL_SULAC
37	40	41.7	99	1	VAT_CAMV
38	40	41.7	101	1	VE7_HPV21
39	40	41.7	102	1	VE7_HPV20
40	40	41.7	104	1	VE7_HPV26
41	40	41.7	111	1	VE7_HPV07
42	40	41.7	111	1	VE7_HPV40
43	40	41.7	159	1	VAT_CAMVC
44	40	41.7	159	1	VAT_CAMVE
45	40	41.7	159	1	VAT_CAMVN
ALIGNMENTS					
OS	Human papillomavirus type 18.				OR8316 treponema p
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;				P22223 homo sapien
OC	Papillomavirus.				P1512 sulfolobus
OX	NCBI_TaxID=10582;				001087 cauliflower
RN	[1]				P50779 human papil
RX	SEQUENCE FROM N.A.				P36824 human papil
RX	MEDLINE=87283882; PubMed=3039146;				P36826 human papil
RA	Cole S.T., Danos O.;				P36829 human papil
RT	"Nucleotide sequence and comparative analysis of the human				P03549 cauliflower
RT	papillomavirus type 18 genome. Phylogeny of papillomaviruses and				002966 cauliflower
RT	repeated structure of the E6 and E7 gene products.";				000965 cauliflower
RL	J. Mol. Biol. 193:599-608(1987).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88188247; PubMed=2833614;				
RA	Inagaki Y., Tsunokawa Y., Takebe N., Nawa H., Nakanishi S.;				
RA	Terada M., Sugimura T.;				
RT	"Nucleotide sequences of cDNAs for human papillomavirus type 18				
RT	transcripts in HeLa cells.";				
RL	J. Virol. 62:1640-1646(1988).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87053870; PubMed=3023067;				
RA	Schneider-Gaedcke A., Schwarz E.;				
RT	"Different human cervical carcinoma cell lines show similar				
RT	transcription patterns of human papillomavirus type 18 early genes.";				
RL	EMBO J. 5:2285-2292(1986).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87218459; PubMed=3034571;				
RA	Seedorf K., Oltersdorf T., Kraemer G., Roewkamp W.;				
RT	"Identification of early proteins of the human papilloma viruses type				
RT	16 (HPV 16) and type 18 (HPV 18) in cervical carcinoma cells.";				
RL	EMBO J. 6:139-144(1987).				
CC	- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING				
CC	ACTIVITIES.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement. (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
CC	EMBL: X05015; CAA28665.1; -				
CC	EMBL: M20324; AAA9513.1; -				

DR EMBL; M20325; AAA95515.1; -;  
 DR EMBL; M26798; AAA46947.1; -;  
 DR EMBL; X04773; CAA28467.1; -;  
 DR EMBL; A06324; CAA00540.1; -;  
 DR EMBL; A06328; CAA00543.1; -;  
 DR PIR; H26251; W7EL18;  
 DR InterPro; IPR00148; Papv1\_E7.  
 DR Pfam; PF00527; E7; 1.  
 KW Early protein; Transcription regulation; Transforming protein;  
 KM DNA-binding; Trans-acting factor.  
 FT DOMAIN 63 66 C-XX-C MOTIF 1.  
 FT CONFLICT 98 101 C-XX-C MOTIF 2.  
 FT CONFLICT 73 73 K -> E (IN REF. 1 AND 4).  
 SQ SEQUENCE 105 AA; 11995 MW; 2CDB119534D0186A CRC64;

Query Match Best Local Similarity 100.0%; Score 96; DB 1; Length 105;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLODIVLHPONEIPV 19  
 DB 5 KATLODIVLHPONEIPV 23

RESULT 2  
 ID VE7\_HPVS5 STANDARD; PRT; 106 AA.  
 AC P21735;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS -Human papillomavirus type 45.  
 OC Viruses; dsDNA viruses; no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 ON NCBI\_TaxID=10593;  
 RX MEDLINE=94265501; PubMed=8205838;  
 RA Delius H., Hofmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types.";  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Kaplan J.B., Burk R.D.;  
 RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
 CC ACTIVITIES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X74479; CAA52574.1; -;  
 DR EMBL; M38198; AAA46974.1; -;  
 DR PIR; S36562; S36562.  
 DR InterPro; IPR00148; Papv1\_E7.  
 DR Pfam; PF00527; E7; 1.  
 KW Early protein; Transcription regulation; Transforming protein;  
 KM DNA-binding; Trans-acting factor.  
 FT DOMAIN 64 67 C-XX-C MOTIF 1.  
 FT CONFLICT 99 102 C-XX-C MOTIF 2.  
 FT CONFLICT 3 3 G -> E (IN REF. 2).  
 FT CONFLICT 5 5 R -> Q (IN REF. 2).  
 SQ SEQUENCE 106 AA; 12049 MW; 4C53808A7285ADA1 CRC64;

Query Match Best Local Similarity 74.5%; Score 71.5; DB 1; Length 106;  
 Matches 15; Conservative 3; Mismatches 1; Indels 1; Caps 1;  
 QY 1 KATLODIVLHPONEIPV 19  
 DB 5 RETLODIVLHPONEIPV 24

RESULT 3  
 ID VE7\_HPVS6 STANDARD; PRT; 105 AA.  
 AC P36833;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 56.  
 OC Viruses; dsDNA viruses; no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 ON NCBI\_TaxID=10596;  
 RX MEDLINE=94265501; PubMed=8205838;  
 RA Delius H., Hofmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types.";  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kaplan J.B., Burk R.D.;  
 RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
 CC ACTIVITIES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X74483; CAA52597.1; -;  
 DR PIR; S36580; S36580.  
 DR InterPro; IPR00148; Papv1\_E7.  
 DR Pfam; PF00527; E7; 1.  
 KW Early protein; Transcription regulation; Transforming protein;  
 KM DNA-binding; Trans-acting factor.  
 FT DOMAIN 65 68 C-XX-C MOTIF 1.  
 FT CONFLICT 98 101 C-XX-C MOTIF 2.  
 FT CONFLICT 5 5 R -> Q (IN REF. 2).  
 SQ SEQUENCE 105 AA; 11892 MW; 35424B643BAE0183 CRC64;

Query Match Best Local Similarity 57.3%; Score 55; DB 1; Length 105;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 TLQDIVLHPONEIPV 17  
 DB 7 TLQDIVLHPONEIPV 21

RESULT 4  
 ID VE7\_HPVS1 STANDARD; PRT; 101 AA.  
 AC P26558;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE E7 PROTEIN.  
 GN E7.  
 OS Human papillomavirus type 51.  
 OC Viruses; dsDNA viruses; no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 ON NCBI\_TaxID=10595;  
 RX MEDLINE=94265501; PubMed=8205838;  
 RA Delius H., Hofmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types.";  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kaplan J.B., Burk R.D.;  
 RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
 CC ACTIVITIES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X74479; CAA52574.1; -;  
 DR EMBL; M38198; AAA46974.1; -;  
 DR PIR; S36562; S36562.  
 DR InterPro; IPR00148; Papv1\_E7.  
 DR Pfam; PF00527; E7; 1.  
 KW Early protein; Transcription regulation; Transforming protein;  
 KM DNA-binding; Trans-acting factor.  
 FT DOMAIN 64 67 C-XX-C MOTIF 1.  
 FT CONFLICT 99 102 C-XX-C MOTIF 2.  
 FT CONFLICT 3 3 G -> E (IN REF. 2).  
 FT CONFLICT 5 5 R -> Q (IN REF. 2).  
 SQ SEQUENCE 106 AA; 12049 MW; 4C53808A7285ADA1 CRC64;



```

RP SEQUENCE FROM N.A.
RX MEDLINE-91303675; PubMed-1649326;
RA Lunqu O., Crum C.P., Silverstein S.J.;
RT "Biologic properties and nucleotide sequence analysis of human
RL papillomavirus type 51."
RT J. Virol. 65:4216-4225(1991).
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62877; -; NOT_ANNOTATED_CDS.
DR PIR; F40415; W7ML51.
DR InterPro; IPR000148; Papv1_E7.
DR Pfam; PF00527; E7; 1.
KW Early protein; Transcription regulation; Transforming protein;
KM DNA-binding; Trans-acting factor.
FT DOMAIN 62 65 C-XX-C MOTIF 1.
FT DOMAIN 95 98 C-XX-C MOTIF 2.
SO SEQUENCE 101 AA; 11339 MW; 65E85C554099801A CRC64;

Query Match 56.28; Score 54; DB 1; Length 101;
Best Local Similarity 71.4%; Pred. No. 0.037;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 LODIVLHPONEI 17
Db 8 LKDVVLHPOTEI 21

RESULT 5
VE7_HPV39 STANDARD; PRT; 109 AA.
AC P24637;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 39.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxId=10588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91135017; PubMed-1847266;
RA Volpers C., Strebeck R.E.;
RT "Genome organization and nucleotide sequence of human papillomavirus
RT type 39."
RT J. Virol. 181:419-423(1991).
RL -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62849; AAA47051.1; -
DR PIR; B38502; W7ML39.
DR InterPro; IPR000148; Papv1_E7.
DR Pfam; PF00527; E7; 1.
KW Early protein; Transcription regulation; Transforming protein;

```

```

KM DNA-binding; Trans-acting factor.
FT DOMAIN 66 69 C-XX-C MOTIF 1.
FT DOMAIN 101 104 C-XX-C MOTIF 2.
SO SEQUENCE 109 AA; 12497 MW; E4743E5A6798C2B CRC64.

```

```

Query Match 54.7%; Score 52.5; DB 1; Length 109;
Best Local Similarity 70.0%; Pred. No. 0.071;
Matches 14; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

```

```

QY 1 KATLQDIVLHPONEI-PV 19
Db 5 KPTLQDIVLHPYNEIQPV 24

```

## RESULT 6

```

VE7_HPV70 STANDARD; PRT; 109 AA.
AC P50785;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 70.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxId=39457;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96249586; PubMed-8815087;
RA Forslund O., Hansson B.G.;
RT "Human papillomavirus type 70 genome cloned from overlapping PCR
RT products: complete nucleotide sequence and genomic organization."
RT J. Clin. Microbiol. 34:802-809(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-97060129; PubMed-8904450;
RA Longuet M., Beaudeau S., Orth G.;
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,
RT related to the potentially oncogenic HPV39."
RT J. Clin. Microbiol. 34:738-744(1996).
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U21941; AAC54851.1; -
DR EMBL; U22461; AAC54881.1; -
DR InterPro; IPR000148; Papv1_E7.
DR Pfam; PF00527; E7; 1.
KW Early protein; Transcription regulation; Transforming protein;
KM DNA-binding; Trans-acting factor.
FT DOMAIN 66 69 C-XX-C MOTIF 1.
FT DOMAIN 101 104 C-XX-C MOTIF 2.
SO SEQUENCE 109 AA; 12657 MW; 2FD3C6A3DFAF691 CRC64;

Query Match 53.6%; Score 51.5; DB 1; Length 109;
Best Local Similarity 65.0%; Pred. No. 0.1;
Matches 13; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

```

```

RESULT 7
VE7_HPV68 STANDARD; PRT; 110 AA.
AC P54668;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 68.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=45240;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97060129; PubMed=8904450;
RA Longuet M., Beaudeau S., Orth G.;
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,
RL J. Clin. Microbiol. 34:738-744(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67160; CAA47633.1;
DR InterPro: IPR000148; Papv1_E7.
DR Pfam: PF00527; E7; 1.
KW Early protein; Transcription regulation; Transforming protein;
KM DNA-binding; Trans-acting factor.
FT DOMAIN 67 70 C-XX-C MOTIF 1.
FT DOMAIN 102 105 C-XX-C MOTIF 2.
SQ SEQUENCE 110 AA; 12756 MW; B52459199048093D CRC64;

Query Match 52.6%; Score 50.5; DB 1; Length 110;
Best Local Similarity 65.0%; Pred. No. 0.16;
Matches 13; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 KATLQDIVLHLEPONEI-pv 19
   1:1:1111111111111
Db 5 KPTVQEVLELCPNIEIPV 24

RESULT 8
VE7_HPV66 STANDARD; PRT; 105 AA.
AC Q80936;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 66.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=37119;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U31794; AAA79500.1;
DR InterPro: IPR000148; Papv1_E7.
DR Pfam: PF00527; E7; 1.
KW Early protein; Transcription regulation; Transforming protein;
KM DNA-binding; Trans-acting factor.
FT DOMAIN 66 69 C-XX-C MOTIF 1.
FT DOMAIN 98 101 C-XX-C MOTIF 2.
SQ SEQUENCE 105 AA; 11911 MW; C09339F42F62AFFA CRC64;

Query Match 52.1%; Score 50; DB 1; Length 105;
Best Local Similarity 60.0%; Pred. No. 0.18;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 TLQDIVLHLEPONEI 17
   1111111111111
Db 7 TLQEVLELAPOTET 21

RESULT 9
VE7_HPV6E STANDARD; PRT; 110 AA.
AC P27963;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type ME180.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374616; PubMed=1716694.
RA Reuter S., Delius H., Kahn T., Hofmann B., Zur Hausen H., Schwarz E.;
RT "Characterization of a novel human papillomavirus DNA in the cervical
RL J. Virol. 65:5564-5568(1991).
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M73258; AAF14012.1;
DR PIR: D40509; W7MLPR.
DR InterPro: IPR000148; Papv1_E7.
DR Pfam: PF00527; E7; 1.
KW Early protein; Transcription regulation; Transforming protein;
KM DNA-binding; Trans-acting factor.
FT DOMAIN 69 72 C-XX-C MOTIF 1.
FT DOMAIN 102 105 C-XX-C MOTIF 2.
SQ SEQUENCE 110 AA; 12672 MW; 90ECC0E004AA666B CRC64;

Query Match 50.5%; Score 48.5; DB 1; Length 110;
Best Local Similarity 65.0%; Pred. No. 0.33;
Matches 13; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 KATLQDIVLHLEPONEI-pv 19
   1:1:1111111111111
Db 5 KPTVQEVLELCPNIEIPV 24

```



```

CC NCBI_TaxID=10570;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=01135018; PubMed=1847267;
CC Ostrow R.S.; Labresh K.V.; Faras A.J.;
CC "Characterization of the complete RHPV 1 genomic sequence and an
CC integration locus from a metastatic tumor.";
CC Virology 181:424-429(1991).
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M60184; AAA/9312.1; -.
CC PIR: B38503; W7WLR1.
CC InterPro: IPR000148; Papv1_E7.
CC Pfam: PF00527; E7; 1.
CC DR Early protein; Transcription regulation; Transforming protein;
CC DNA-binding; Trans-acting factor.
CC FT DOMAIN 73 76 C-XX-C MOTIF 1;
CC 106 109 C-XX-C MOTIF 2.
CC SQ SEQUENCE 113 AA; 12818 MW; 9C404BDAC1298306 CRC64;

OY 1 KATLQDIVLHLEP 13
DB 5 KPLEDIVLDLDP 17

Query Match 46.9%; Score 45; DB 1; Length 113;
Best Local Similarity 69.2%; Pred. No. 1.3;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KATLQDIVLHLEP 13
DB 5 KPLEDIVLDLDP 17

RESULT 13
SOK1_YEAST STANDARD: PRT; 901 AA.
AC PA0337;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE SOK1 PROTEIN.
GN SOK1 OR YDR006C OR YD8119.12C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycs.
CC NCBI_TaxID=4932;
CC [1]
CC SEQUENCE OF 337-901 FROM N.A.
CC STRAIN=SGY101;
CC MEDLINE=94344070; PubMed=8065298;
CC Ward M.P.; Garrett S.;
CC "Suppression of a yeast cyclic AMP-dependent protein kinase defect by
CC overexpression of SOK1, a yeast gene exhibiting sequence similarity
CC to a developmentally regulated mouse gene.";
CC Mol. Cell. Biol. 14:5619-5627(1994).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=S288C / AB972;
CC Murphy L.; Richards C.; Gentles S.; Harris D.; Barrell B.G.;
CC Rajandream M.A.;
CC Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HIGH COPY SUPPRESSOR OF A CYCLIC AMP-DEPENDENT
CC PROTEIN KINASE MUTANT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: SOME TO MOUSE TESTIS-SPECIFIC PROTEIN PPS13.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U12027; AAA56922.1; -.
CC DR EMBL: Z48008; CAA88066.1; -.
CC DR PIR: S47913; S47913.
CC DR SGD: S0002413; SOK1.
CC KW Nuclear protein.
CC FT CONFLICT 645 A -> R (IN REF. 1).
CC SQ SEQUENCE 901 AA; 101081 MW; C900DF289541B98 CRC64;

OY 3 TLQDIVLHLEPQNPV 19
DB 452 TLQDIVLHLEPQNPV 468

Query Match 46.9%; Score 45; DB 1; Length 901;
Best Local Similarity 47.1%; Pred. No. 13;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 3 TLQDIVLHLEPQNPV 19
DB 452 TLQDIVLHLEPQNPV 468

RESULT 14
VE7_HPVS4 STANDARD: PRT; 95 AA.
AC 081019;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human papillomavirus type 54.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
CC NCBI_TaxID=37113;
CC [1]
CC SEQUENCE FROM N.A.
CC RA Delius H.;
CC Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U37488; AAA79188.1; -.
CC DR InterPro: IPR000148; Papv1_E7.
CC DR Pfam: PF00527; E7; 1.
CC KW Early protein; Transcription regulation; Transforming protein;
CC DNA-binding; Trans-acting factor.
CC FT DOMAIN 88 91 C-XX-C MOTIF.
CC SQ SEQUENCE 95 AA; 10565 MW; 6A543269F2561891 CRC64;

OY 2 ATLQDIVLHLEPQ 14
DB 6 ATLQDIVLHLEPQ 18

Query Match 45.8%; Score 44; DB 1; Length 95;
Best Local Similarity 61.5%; Pred. No. 1.6;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ATLQDIVLHLEPQ 14
DB 6 ATLQDIVLHLEPQ 18

RESULT 15
VE7_HPVS4

```

```

ID VE7_HPV24 STANDARD; PRT: 96 AA.
AC P50782;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 01-OCT-1996 (Rel. 34, last annotation update)
DE E7 PROTEIN.
GN E7.
OS Human Papillomavirus type 24.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID-37956;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/Genbank/DDJ databases.
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U31782; AAA79416.1; -;
DR InterPro: IPR000148; Papv1_E7.
DR Pfam: PF00527; E7; 1.
KW Early protein; Transcription regulation; Transforming protein;
KW DNA-binding; Trans-acting factor.
FT DOMAIN 83 86 C-XX-C MOTIF.
SQ SEQUENCE 96 AA; 10710 MW; BDE40A581E0829C5 CRC64;

```

```

Query Match 45.8%; Score 44; DB 1; Length 96;
Best Local Similarity 50.0%; Pred. No. 1.6;
Matches 13; Conservative 2; Mismatches 3; Indels 8; Gaps 2;

```

```

Qy . 1 KATLDDIVLHL-EPQ-----NEIP 18
      : ||||| | | | | | | | | | |
Db 5 EVTLDDIVLELTPQTVDLHCEELP 30

```

Search completed: February 13, 2002, 10:02:21  
Job time: 250 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 10:00:44 : Search time 39.34 Seconds

(without alignments)  
36.790 Million cell updates/sec

Title: US-09-664-225-154

Sequence: 1 KATLQDIVLHLEPONEIPV 19

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR\_68:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	105	1 W7WL18	E7 protein - human
2	71.5	74.5	106	2 S36562	E7 protein - human
3	55	57.3	105	3 S36580	E7 protein - human
4	54	56.2	101	1 W7WL51	E7 protein - human
5	52.5	54.7	109	1 W7WL39	E7 protein - human
6	50	52.1	105	2 B44890	E7 protein - human
7	48.5	50.5	110	1 W7WLP8	E7 protein - human
8	47	49.0	369	2 T29120	hypothetical prote
9	47	49.0	803	1 R6B1W6	regulatory protein
10	46	47.9	249	2 T39606	40S ribosomal prot
11	45	46.9	113	1 W7WLR1	E7 protein - rhesu
12	45	45.3	901	2 S50987	SOK1 protein - yea
13	44	45.8	97	2 S36516	E7 protein - human
14	44	45.8	99	1 W7WLR3	E7 protein - human
15	44	45.8	1616	2 T16600	vitellinogenin vit-1
16	43.5	45.3	462	2 T25544	hypothetical prote
17	43	44.8	99	2 S36574	E7 protein - human
18	43	44.8	1613	2 A43081	vitellinogenin vit-2
19	42	43.8	98	1 W7WLR1	E7 protein - human
20	42	43.8	98	1 W7WLR6	E7 protein - human
21	42	43.8	288	2 H81326	probable signal re
22	42	43.8	394	2 A72223	panthothenate metab
23	42	43.8	490	2 A65191	hypothetical 54.1
24	42	43.8	490	2 F86073	probable GMP-bindi
25	42	43.8	634	1 B32392	protein kinase C (
26	42	43.8	4436	2 E71086	hypothetical prote
27	42	43.8	5032	1 A35041	ryanodine receptor
28	41.5	43.2	171	2 H65048	hypothetical prote
29	41.5	43.2	171	2 H85916	hypothetical prote

30	41	42.7	98	1 W7WL31	E7 protein - human
31	41	42.7	126	2 S09919	hypothetical prote
32	41	42.7	132	2 S77000	hypothetical prote
33	41	42.7	187	2 C71317	hypothetical prote
34	41	42.7	211	2 A83724	hypothetical prote
35	41	42.7	260	2 S77107	hypothetical prote
36	41	42.7	264	2 S74686	hypothetical prote
37	41	42.7	268	2 S75033	hypothetical prote
38	41	42.7	290	2 H83286	probable chemotaxi
39	41	42.7	429	1 IJH0CP	cadherin 3 precurs
40	41	42.7	845	2 T00071	hypothetical prote
41	41	42.7	880	1 B33926	DNA-directed RNA p
42	41	42.7	2037	2 T16881	hypothetical prote
43	41	42.7	5369	2 T44807	mycosubtilin synth
44	40	41.7	104	2 S36545	E7 protein - human
45	40	41.7	111	2 S36585	E7 protein - human

#### ALIGNMENTS

```

RESULT 1
W7WL18
E7 protein - human papillomavirus type 18
C:Species: human papillomavirus type 18
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: B26165; H26251
R:Seedorf, K.; Oltersdorf, T.; Kraemer, G.; Roewekamp, W.
EMBO J. 6, 139-144, 1987
A>Title: Identification of early proteins of the human papilloma viruses type 16 (HPV
A:Reference number: A91068; MWID:87218459
A:Accession: B26165
A:Molecule type: DNA
A:Residues: 1-105 <SEE>
A:Cross-references: GB:X04773; NID:960876; PIDN:CAA28467.1; PID:960878
R:Cooley, S.F.; Danos, O.
J. Mol. Biol. 193, 599-608, 1987
A>Title: Nucleotide sequence and comparative analysis of the human papillomavirus typ
A:Reference number: A92937; MWID:87283882
A:Accession: H26251
A:Molecule type: DNA
A:Residues: 1-105 <COL>
A:Cross-references: GB:X05015; NID:960975; PIDN:CAA28665.1; PID:960977
C:Superfamily: papillomavirus E7 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 100.0%; Score 96; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLQDIVLHLEPONEIPV 19
   |||||
Db 5 KATLQDIVLHLEPONEIPV 23

RESULT 2
S36562
E7 protein - human papillomavirus type 45
C:Species: human papillomavirus type 45
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36562
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36562
A:Molecule type: DNA
A:Residues: 1-106 <DEL>
A:Cross-references: EMBL:X74479; NID:g397022; PIDN:CAA52574.1; PID:g397024
C:Superfamily: papillomavirus E7 protein
C:Keywords: DNA binding; early protein; transcription regulation

```

Query Match 74.5%; Score 71.5; DB 2; Length 106;  
Best Local Similarity 75.0%; Pred. No. 0.00018;  
Matches 15; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

OY 1 KATLODIVLHLEPONEI-pv 19  
: |||:|||||: ||  
Db 5 RETLOEIVLHLEPONEIDPV 24

## RESULT 3

S36580  
E7 protein - human papillomavirus type 56  
C:Species: human papillomavirus type 56  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
A:Accession: S36580  
R:Delius, H.; Hofmann, B.  
submitted to the EMBL Data Library, August 1993  
A:Description: Primer-directed sequencing of human papillomavirus types.  
A:Reference number: S36469  
A:Accession: S36580  
A:Molecule type: DNA  
A:Residues: 1-105 <DEL>  
A:Cross-references: EMBL:X74483; NID:g397053; PIDN:CAA52597.1; PID:g397055  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 57.3%; Score 55; DB 2; Length 105;  
Best Local Similarity 73.3%; Pred. No. 0.084;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 TLQDIVLHLEPONEI 17  
|||:| | | | |  
Db 7 TLQDIVLHLEPOTEI 21

## RESULT 4

W7WL51  
E7 protein - human papillomavirus type 51  
C:Species: human papillomavirus type 51  
A:Note: host Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: F40415  
R:Lungu, O.; Crum, C.P.; Silverstein, S.J.  
J. Virol. 65, 4216-4225, 1991  
A:Title: Biologic properties and nucleotide sequence analysis of human papillomavirus type  
A:Reference number: A40415; MUID:91303675  
A:Accession: F40415  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-101 <LUN>  
A:Cross-references: GB:M62877  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 56.2%; Score 54; DB 1; Length 101;  
Best Local Similarity 71.4%; Pred. No. 0.12;  
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 LQDIVLHLEPONEI 17  
|:|:| | | | |  
Db 8 LKDVVLEHLPOTEI 21

## RESULT 5

W7WL39  
E7 protein - human papillomavirus type 39  
C:Species: human papillomavirus type 39  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999  
A:Accession: B38502

R:Volpers, C.; Strebeck, R.E.  
Virology 181, 419-423, 1991  
A:Title: Genome organization and nucleotide sequence of human papillomavirus type 39.  
A:Reference number: A38502; MUID:91135017  
A:Accession: B38502  
A:Status: translation not shown

A:Molecule type: DNA  
A:Residues: 1-109 <VOI>  
A:Cross-references: GB:M62849; EMBL:M38185; NID:g333245; PIDN:AAA47051.1; PID:g463187  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 54.7%; Score 52.5; DB 1; Length 109;  
Best Local Similarity 70.0%; Pred. No. 0.22;  
Matches 14; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 KATLODIVLHLEPONEI-pv 19  
| | | : | | | | |  
Db 5 KPTLOEIVLDCPYNEIQPV 24

## RESULT 6

B44890  
E7 protein - human papillomavirus type 66  
C:Species: human papillomavirus type 66  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 22-May-1998  
C:Accession: B44890  
R:Tawheed, A.R.; Beaudenon, S.; Favre, M.; Orth, G.  
J. Clin. Microbiol. 29, 2656-2660, 1991  
A:Title: Characterization of human papillomavirus type 66 from an invasive carcinoma  
A:Reference number: A44890; MUID:92129556  
A:Accession: B44890  
A:Molecule type: DNA  
A:Residues: 1-105 <TAM>  
A:Note: sequence extracted from NCBI backbone (NCBIN:78637, NCBI:P:78639)  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 52.1%; Score 50; DB 2; Length 105;  
Best Local Similarity 60.0%; Pred. No. 0.54;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 TLQDIVLHLEPONEI 17  
|||:| | | | |  
Db 7 TLQEIVLELPOTEI 21

## RESULT 7

W7WLPR  
E7 protein - human papillomavirus type ME180 (provirus)  
C:Species: human papillomavirus type ME180  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 28-Jul-2000  
C:Accession: D40509  
R:Reuter, S.; Delius, H.; Kahn, T.; Hofmann, B.; zur Hausen, H.; Schwartz, E.  
J. Virol. 65, 5564-5568, 1991  
A:Title: Characterization of a novel human papillomavirus DNA in the cervical carcinoma  
A:Reference number: A40509; MUID:91374616  
A:Accession: D40509  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-110 <REV>  
A:Cross-references: GB:M73258; NID:g184383; PIDN:AAF14012.1; PID:g6478872  
C:Superfamily: papillomavirus E7 protein  
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 50.5%; Score 48.5; DB 1; Length 110;  
Best Local Similarity 65.0%; Pred. No. 0.99;  
Matches 13; Conservative 2; Mismatches 4; Indels 1; Gaps 1;



OY 1 KATLQDIVLHLEPONEIPV 19  
 | | | | | | | | | |  
 DB 5 KPTVQETIVLELCPNRIEIPV 24

## RESULT 8

hypothetical protein - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999  
 C:Accession: T29120  
 R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, August 1998  
 A:Reference number: Z17215  
 A:Accession: T29120  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-369 <P&R>  
 A:Cross-references: EMBL:AL031350; NID:e1316892; PID:e1316903; PIDN:CAA20502.1  
 C:Genetics:  
 A:Note: SCLF2.11

Query Match 49.0%; Score 47; DB 2; Length 369;  
 Best Local Similarity 61.1%; Pred. No. 6.7;  
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 ATLQDIVLHLEPONEIPV 19  
 | | | | | | | | | |  
 DB 246 ATYQGIHRLPPRAEIPV 263

## RESULT 9

RGBYW6  
 regulatory protein SM16 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein I9470.8; protein YLR182w  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 12-Nov-1999  
 C:Accession: S03161; S51427  
 R:Bredeno, L.; Masnyth, K.  
 Nature 329, 651-654, 1987  
 A:Title: Similarity between cell-cycle genes of budding yeast and fission yeast and the  
 A:Reference number: S03161; MUID:88014241  
 A:Accession: S03161  
 A:Molecule type: DNA  
 A:Residues: 1-803 <BR>  
 A:Cross-references: EMBL:X06238; NID:94597; PIDN:CAA29581.1; PID:94598  
 R:Wohlmann, P.  
 submitted to the EMBL Data Library, November 1994  
 A:Description: The sequence of S. cerevisiae cosmid 9470.  
 A:Reference number: S51414  
 A:Accession: S51427  
 A:Molecule type: DNA  
 A:Residues: 1-803 <WOH>  
 A:Cross-references: EMBL:U17246; NID:9577192; PIDN:AAB67460.1; PID:9577200; GSPDB:GN0001  
 C:Genetics:  
 A:Gene: SGD:SM16; SD511; PSTL: MIPS:YLR182w  
 A:Cross-references: SGD:S0004172; MIPS:YLR182w  
 A:Map position: 12R  
 C:Superfamily: regulatory protein SM16; ankyrin repeat homology  
 C:Keywords: DNA binding; transcription regulation  
 F:317-349/Domain: ankyrin repeat homology <ANI>  
 F:469-501/Domain: ankyrin repeat homology <AN2>

Query Match 49.0%; Score 47; DB 1; Length 803;  
 Best Local Similarity 50.0%; Pred. No. 16;  
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 4 LODIVLHLEPONEIPV 19  
 | | | | | | | | | |  
 DB 3 LEEVYATLCPHNEIPL 18

RESULT 10

40s ribosomal protein s3 - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
 C:Accession: T39606  
 R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.  
 submitted to the EMBL Data Library, May 1998  
 A:Reference number: Z21866  
 A:Accession: T39606

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-249 <LYN>  
 A:Cross-references: EMBL:AL023554; PIDN:CAA19033.1; GSPDB:GN00067; SPDB:SPBC1665.14C  
 A:Experimental source: strain 97zh-; cosmid c1665  
 C:Genetics:  
 A:Gene: SPDB:SPBC1665.14C  
 A:Map position: 2  
 A:Introns: 12/3  
 C:Superfamily: rat ribosomal protein S3

Query Match 47.9%; Score 46; DB 2; Length 249;  
 Best Local Similarity 47.4%; Pred. No. 6.2;  
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 KATLQDIVLHLEPONEIPV 19  
 | | | | | | | | | |  
 DB 200 KKSIPDIVVVDPKREEPI 218

## RESULT 11

W7MR1  
 E7 protein - rhesus papillomavirus (type 1)  
 C:Species: rhesus papillomavirus  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Feb-1997  
 C:Accession: B38503  
 R:Ostrow, R.S.; Labresh, K.V.; Farris, A.J.  
 Virology 181, 424-429, 1991  
 A:Title: Characterization of the complete RHPV 1 genomic sequence and an integration  
 A:Reference number: B38503; MUID:91135018  
 A:Accession: B38503  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-113 <OST>  
 A:Cross-references: EMBL:M37717  
 C:Superfamily: papillomavirus E7 protein  
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 45.9%; Score 45; DB 1; Length 113;  
 Best Local Similarity 69.2%; Pred. No. 3.8;  
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KATLQDIVLHLEP 13  
 | | | | | | | | | |  
 DB 5 KPTLEDIVLDLP 17

## RESULT 12

S50987  
 SOK1 protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein Y08119.12c; protein Y08006c; suppressor of kinase  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 11-Feb-1995 #sequence\_revision 11-Aug-1995 #text\_change 21-Jul-2000  
 C:Accession: S50987; S47913; A56237  
 R:Murphy, L.; Richards, C.; Gentles, S.; Harris, D.  
 submitted to the EMBL Data Library, January 1995  
 A:Reference number: S50976  
 A:Accession: S50987  
 A:Molecule type: DNA  
 A:Residues: 1-901 <MUR>



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 13, 2002, 09:59:58 ; Search time 32.4 seconds  
(without alignments)  
13.196 Million cell updates/sec

Title: US-09-664-225-154

Perfect score: 96

Sequence: 1 KATLQDIVLHPEQNEIPV 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents-AA:\*

1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*

2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*

3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*

4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*

5: /cgn2\_6/prodata/2/1aa/6CTUS.COMB.pep:\*

6: /cgn2\_6/prodata/2/1aa/backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	30	2	US-08-934-915-61 Sequence 61, Appl
2	96	100.0	272	1	US-08-117-083-13 Sequence 13, Appl
3	63	65.6	48	1	US-08-466-285-5 Sequence 5, Appl
4	46	47.9	23	1	US-08-188-582-27 Sequence 27, Appl
5	46	47.9	23	1	US-08-646-715-27 Sequence 27, Appl
6	44	45.8	9	4	US-09-124-671-10 Sequence 10, Appl
7	42	43.8	601	2	US-08-606-288-7 Sequence 7, Appl
8	42	43.8	601	2	US-08-606-288-10 Sequence 10, Appl
9	42	43.8	601	3	US-09-347-483-7 Sequence 7, Appl
10	42	43.8	601	3	US-09-347-483-10 Sequence 10, Appl
11	40	41.7	342	1	US-08-499-215-2 Sequence 2, Appl
12	39	40.6	516	4	US-08-379-580-4 Sequence 4, Appl
13	39	40.6	870	1	US-08-190-687B-25 Sequence 25, Appl
14	39	40.6	870	5	PCr-US94-00198-5 Sequence 5, Appl
15	39	40.6	1005	2	US-08-935-450-2 Sequence 2, Appl
16	39	40.6	1047	1	US-08-190-687B-8 Sequence 8, Appl
17	38	39.6	30	2	US-08-934-915-62 Sequence 62, Appl
18	38	39.6	377	2	US-08-839-581A-31 Sequence 31, Appl
19	38	39.6	377	4	US-09-023-591A-31 Sequence 31, Appl
20	38	39.6	496	3	US-08-926-842B-12 Sequence 12, Appl
21	38	39.6	518	4	US-09-240-915-8 Sequence 8, Appl
22	38	39.6	518	4	US-09-591-435-8 Sequence 8, Appl
23	38	39.6	547	1	US-08-314-615-1 Sequence 1, Appl
24	38	39.6	547	1	US-08-314-615-1 Sequence 1, Appl
25	38	39.6	547	1	US-08-473-981A-6 Sequence 6, Appl
26	38	39.6	547	1	US-08-433-010-1 Sequence 1, Appl
27	38	39.6	547	1	US-08-482-882-1 Sequence 1, Appl

28	38	39.6	547	2	US-08-483-389-1 Sequence 1, Appl
29	38	39.6	547	2	US-08-487-113D-1 Sequence 1, Appl
30	38	39.6	547	2	US-08-473-503-1 Sequence 1, Appl
31	38	39.6	547	2	US-08-483-932-1 Sequence 1, Appl
32	38	39.6	547	2	US-08-474-087-6 Sequence 6, Appl
33	38	39.6	547	2	US-08-720-420A-1 Sequence 1, Appl
34	38	39.6	547	3	US-08-714-017-1 Sequence 1, Appl
35	38	39.6	547	3	US-08-863-790-1 Sequence 1, Appl
36	38	39.6	547	3	US-08-475-680-1 Sequence 1, Appl
37	38	39.6	547	4	US-08-296-749-1 Sequence 1, Appl
38	38	39.6	1311	1	US-08-340-011-5 Sequence 5, Appl
39	38	39.6	1311	3	US-08-901-710-5 Sequence 5, Appl
40	38	39.6	1338	3	US-08-750-141A-3 Sequence 3, Appl
41	38	39.6	4654	4	US-08-476-515A-84 Sequence 84, Appl
42	38	39.6	4655	4	US-08-652-877-84 Sequence 84, Appl
43	38	39.6	4655	4	US-08-652-877-86 Sequence 86, Appl
44	38	39.6	4655	4	US-08-652-877-88 Sequence 88, Appl
45	38	39.6	4655	4	US-08-652-877-90 Sequence 90, Appl

## ALIGNMENTS

RESULT 1

US-08-934-915-61

: Sequence 61, Application US/08934915

: Patent No. 5932412

: GENERAL INFORMATION:

: APPLICANT: DILLNER, JOAKIM

: APPLICANT: DILLNER, LENA

: APPLICANT: CHENG, HWEI-MING

: TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN

: TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,

: TITLE OF INVENTION: USEFUL IN IMMUNOSSAY FOR

: TITLE OF INVENTION: DIAGNOSTIC PURPOSES

: NUMBER OF SEQUENCES: 193

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: MASON & ASSOCIATES, P.A.

: STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500

: CITY: CLEARWATER

: STATE: FLORIDA

: COUNTRY: U.S.A.

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: Windows 3.0

: SOFTWARE: Microsoft Word 6.0

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/934,915

: FILING DATE: 22-SEP-1997

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 07/949,836

: FILING DATE:

: ATTORNEY/AGENT INFORMATION:

: NAME: LOUISE A. FOUTCH

: REGISTRATION NUMBER: 37,133

: REFERENCE/DOCKET NUMBER: 1946.6

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 813-538-3800

: TELEFAX: 813-538-3820

: TELEX:

: INFORMATION FOR SEQ ID NO: 61:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 30 amino acids

: TYPE: amino acid

: TOPOLOGY: linear

: MOLECULE TYPE: peptide

: US-08-934-915-61

Query Match 100.0%; Score 96; DB 2; Length 30;

Best Local Similarity 100.0%; Pred. No. 9.8e-10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLQDIVLHLEPONEIPV 19  
Db 4 KATLQDIVLHLEPONEIPV 22

## RESULT 2

US-08-117-083-13

; Sequence 13, Application US/08117083  
; Patent No. 5719054  
; GENERAL INFORMATION:  
; APPLICANT: Boursnell, Michael E.  
; APPLICANT: Inglis, Stephen C.  
; APPLICANT: Munro, Alan J.  
; TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human  
; TITLE OF INVENTION: Papilloma Virus Proteins  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Walter H. Dreger  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/117,083  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Walter H.  
; REGISTRATION NUMBER: 24,190  
; REFERENCE/DOCKET NUMBER: A-58783  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..272  
; OTHER INFORMATION: /note="Xaa refers to stop codon in  
; OTHER INFORMATION: the open reading frame."  
US-08-117-083-13

Query Match 100.0%; Score 96; DB 1; Length 272;

Best Local Similarity 100.0%; Pred. No. 1.3e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLQDIVLHLEPONEIPV 19  
Db 168 KATLQDIVLHLEPONEIPV 186

## RESULT 3

US-08-466-285-5

; Sequence 5, Application US/08466285  
; Patent No. 5753233  
; GENERAL INFORMATION:  
; APPLICANT: Bleul, Conrad

; APPLICANT: Glissmann, Lutz  
; APPLICANT: Muller, Martin  
; TITLE OF INVENTION: Seroreactive Epitopes On Proteins Of  
; TITLE OF INVENTION: Human Papillomavirus (HPV)18  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,285  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/164,768  
; FILING DATE: 10-DEC-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/947,992  
; FILING DATE: 21-SEP-1992  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/696,953  
; FILING DATE: 08-MAY-1991  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: P 40 15 044.5  
; FILING DATE: 10-MAY-1990  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Manspelzer, David A.  
; REGISTRATION NUMBER: 37,540  
; REFERENCE/DOCKET NUMBER: 05552.1075-03000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)408-4000  
; TELEFAX: (202)408-4400  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 48 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-466-285-5

Query Match 65.6%; Score 63; DB 1; Length 48;

Best Local Similarity 100.0%; Pred. No. 0.00058;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VHLHPONEIPV 19  
Db 1 VHLHPONEIPV 12

## RESULT 4

US-08-188-582-27

; Sequence 27, Application US/08188582  
; Patent No. 5534410  
; GENERAL INFORMATION:  
; APPLICANT: Tjian, Robert  
; APPLICANT: Comai, Lucio  
; APPLICANT: Dynlacht, Brian D.  
; APPLICANT: Hoey, Timothy

APPLICANT: Ruppert, Siegfried  
APPLICANT: Tanese, Naoko  
APPLICANT: Wang, Edith  
APPLICANT: Weinzierl, Robert O.J.  
TITLE OF INVENTION: NUCLEIC ACIDS BINDING PROTEIN ASSOCIATED FACTORS  
TITLE OF INVENTION: NUCLEIC ACIDS BINDING PROTEIN ASSOCIATED FACTORS  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESS: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,582  
FILING DATE: 28-JAN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57650-2/AJY/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-188-582-27

Query Match 47.9%; Score 46; DB 1; Length 23;  
Best Local Similarity 69.2%; Pred. No. 0.17;  
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 6 DIVLHLEPONEIP 18  
I: ||||| || ||  
DB 1 DVOLHLEPONEIP 13  
RESULT 5  
US-08-646-715-27  
Sequence 27, Application US/08646715  
Patent No. 5637686  
GENERAL INFORMATION:  
APPLICANT: Tjian, Robert  
APPLICANT: Comai, Lucio  
APPLICANT: Dynlacht, Brian D.  
APPLICANT: Hoey, Timothy  
APPLICANT: Ruppert, Siegfried  
APPLICANT: Tanese, Naoko  
APPLICANT: Wang, Edith  
APPLICANT: Weinzierl, Robert O.J.  
TITLE OF INVENTION: NUCLEIC ACIDS BINDING PROTEIN ASSOCIATED FACTORS  
TITLE OF INVENTION: NUCLEIC ACIDS BINDING PROTEIN ASSOCIATED FACTORS  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESS: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,715  
FILING DATE: 09-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/188,582  
FILING DATE: 28-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57650-2/AJY/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-646-715-27

Query Match 47.9%; Score 46; DB 1; Length 23;  
Best Local Similarity 69.2%; Pred. No. 0.17;  
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 DIVLHLEPONEIP 18  
I: ||||| || ||  
DB 1 DVOLHLEPONEIP 13

RESULT 6  
US-09-124-671-10  
Sequence 10, Application US/09124671A  
Patent No. 6160088  
GENERAL INFORMATION:  
APPLICANT: Rothman, James  
APPLICANT: Mayhew, Mark  
APPLICANT: Hoe, Mee  
TITLE OF INVENTION: KDE1 RECEPTOR INHIBITORS  
FILE REFERENCE: 31488  
CURRENT APPLICATION NUMBER: US/09/124,671A  
CURRENT FILING DATE: 1998-07-29  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 9  
TYPE: PRT  
ORGANISM: papillomavirus  
US-09-124-671-10

Query Match 45.8%; Score 44; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TLDDIVLHL 11  
I: ||||| || ||  
DB 1 TLDDIVLHL 9

RESULT 7  
US-08-606-288-7  
Sequence 7, Application US/08606288  
Patent No. 5955087

GENERAL INFORMATION:  
APPLICANT: Whittle, N.R.  
APPLICANT: Carmichael, J.P.  
APPLICANT: Connor, S.E.  
APPLICANT: Thompson, H.S.G.  
APPLICANT: Wilson, M.J.  
TITLE OF INVENTION: Polypeptides Useful as Immunotherapeutic  
TITLE OF INVENTION: Agents, and Methods of Polypeptide Preparation  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Suite 3400, Four Embarcadero Center  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/606,288  
FILING DATE: 23-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9503786.7  
FILING DATE: 24-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000034  
FILING DATE: 08-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9515478.7  
FILING DATE: 28-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Walter H. Dreger  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-63284/RMD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 601 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-606-288-7

Query Match 43.8%; Score 42; DB 2; Length 601;  
Best Local Similarity 53.3%; Pred. No. 36;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 3 TLQDIVLHLEPONEI 17  
Db 499 TLKDIVLDLQPPDV 513  
RESULT 8  
US-08-606-288-10  
Sequence 10, Application US/08606288  
Patent No. 5955087  
GENERAL INFORMATION:  
APPLICANT: Whittle, N.R.  
APPLICANT: Carmichael, J.P.  
APPLICANT: Connor, S.E.  
APPLICANT: Thompson, H.S.G.  
APPLICANT: Wilson, M.J.  
TITLE OF INVENTION: Polypeptides Useful as Immunotherapeutic  
TITLE OF INVENTION: Agents, and Methods of Polypeptide Preparation  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Suite 3400, Four Embarcadero Center

CITY: San Francisco  
STATE: California  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/606,288  
FILING DATE: 23-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9503786.7  
FILING DATE: 24-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000034  
FILING DATE: 08-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9515478.7  
FILING DATE: 28-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Walter H. Dreger  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-63284/RMD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 601 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-606-288-10

Query Match 43.8%; Score 42; DB 2; Length 601;  
Best Local Similarity 53.3%; Pred. No. 36;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 3 TLQDIVLHLEPONEI 17  
Db 499 TLKDIVLDLQPPDV 513  
RESULT 9  
US-09-347-483-7  
Sequence 7, Application US/09347483  
Patent No. 6123948  
GENERAL INFORMATION:  
APPLICANT: Whittle, N.R.  
APPLICANT: Carmichael, J.P.  
APPLICANT: Connor, S.E.  
APPLICANT: Thompson, H.S.G.  
APPLICANT: Wilson, M.J.  
TITLE OF INVENTION: Polypeptides Useful as Immunotherapeutic  
TITLE OF INVENTION: Agents, and Methods of Polypeptide Preparation  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Suite 3400, Four Embarcadero Center  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/347,483

Query Match	43.8%	Score 42	DB 3	Length 601
Best Local Similarity	53.3%	Pred. NO. 36		
Matches	8	Conservative	4	Mismatches 3
				Indels 0
				Gaps 0
QY	3	TLQDVLVHLEPQNEL	17	
				:
Db	499	TLKDVITLQDPEDPV	513	

```
QY      4 LODIVLHLEPQNEIP 18
          ||||| ||:  ||
Db      166 LODIVHNLHDLQESP 180
```

```

RESULT 12
US-08-379-580-4
: Sequence 4, Application US/08379580
: Patent No. 6180362
: GENERAL INFORMATION:
: APPLICANT: Duchesne, Marc
: APPLICANT: Schuesignoffer, Fabien
: APPLICANT: Tocque, Bruno
: TITLE OF INVENTION: Peptides Inhibiting Ras Protein
: TITLE OF INVENTION: Activity, Preparation and Use Thereof
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Rhone-Poulenc Rorer Inc.
: STREET: 500 Arcoia Rd. 3C43
: CITY: Collegeville
: STATE: PA
: COUNTRY: USA
: ZIP: 19002
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/379,580
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/FR93/00772
: FILING DATE: 28-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 92/09433
: FILING DATE: 03-JUL-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Smith, Julie K.
: REGISTRATION NUMBER: 38,619
: REFERENCE/DOCKET NUMBER: SP92049-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (610)454-3839
: TELEFAX: (610)454-3808
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 516 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-379-580-4

Query Match 40.6%; Score 39; DB 4; Length 516;
Best Local Similarity 37.5%; Pred. No. 97;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0.

QY 4 LQDIVLHLEPONEIPV 19
: :|||:| :||
Db 463 VSSLVHTEHAHKLPV 478

RESULT 13
US-08-190-687B-25
: Sequence 25, Application US/08190687B
: Patent No. 5760203
: GENERAL INFORMATION:
: APPLICANT: Wong, Gail L.
: APPLICANT: Martin, George
: APPLICANT: McCormick, Francis P.
: APPLICANT: Rubinfeld, Bonnee
: APPLICANT: O'Rourke, Edward C.
: APPLICANT: Clark, Robin
: TITLE OF INVENTION: GAP Gene Sequences
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum

```

```

1 STREET 6300 Sears Tower, 233 South Wacker Drive
2 CITY: Chicago
3 STATE: Illinois
4 COUNTRY: United States of America
5 ZIP: 60606-6402
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: Patent Release #1.0, Version #1.30
12
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/08/190,687B
15 FILING DATE: 02-FEB-1994
16 CLASSIFICATION: 435
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: 07/774,644
19 FILING DATE: 11-OCT-1991
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: 07/260,807
23 FILING DATE: 21-OCT-1988
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: 07/230,761
26 FILING DATE: 10-AUG-1988
27 ATTORNEY/AGENT INFORMATION:
28 NAME: GASS, David A.
29 REGISTRATION NUMBER: 38,153
30 REFERENCE/DOCKET NUMBER: 27527/31898
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: 312/474-6300
33 TELEFAX: 312/474-0448
34
35 TELEX: 25-3856
36
37 INFORMATION FOR SEQ ID NO: 25:
38
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 870 amino acids
41 TYPE: amino acid
42 TOPOLOGY: linear
43
44 MOLECULE TYPE: protein
45
46 US-08-190-687B-25
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
```



APPLICATION NUMBER: US 08/004,824  
 FILING DATE: 15-JAN-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lunn, Paul G.  
 REGISTRATION NUMBER: 32,743  
 REFERENCE/DOCKET NUMBER: DX0352 PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (201)822-7255  
 TELEFAX: (201)822-7039  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 870 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 PCT-US94-00198-5

Query Match 40.6%; Score 39; DB 5; Length 870;  
 Best Local Similarity 37.5%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 4 LODIVLHLEPONEIPV 19  
 : ||||| :  
 Db 416 VSSLVHIEFAHKLPV 431

RESULT 15  
 US-08-935-450-2  
 Sequence 2, Application US/08935450  
 Patent No. 5977311  
 GENERAL INFORMATION:  
 APPLICANT: Nandabalan, Krishnan  
 APPLICANT: Yang, Meijia  
 APPLICANT: Schulz, Vincent  
 TITLE OF INVENTION: 53BP2 COMPLEXES  
 FILE REFERENCE: 7934-054  
 CURRENT APPLICATION NUMBER: US/08/935,450  
 CURRENT FILING DATE: 1997-09-23  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 1005  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-08-935-450-2

Query Match 40.6%; Score 39; DB 2; Length 1005;  
 Best Local Similarity 42.9%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 LODIVLHLEPONEI 17  
 : ||||| :  
 Db 664 IONRYLHVEPEKEV 677

Search completed: February 13, 2002, 09:59:59  
 Job time: 348 sec

**THIS PAGE BLANK (USPTO)**





PI Hedley ML, Urban RC, Chicz RM:  
 DR WPI, 2001-265996/27.  
 XX  
 PT Novel nucleic acids encoding polypeptide polypeptides containing  
 PT multiple epitopes from one or more proteins, useful for treating tumors  
 PT and as vaccines against pathogenic agents  
 XX  
 PS Disclosure; Fig 5; 64pp; English.

CC This invention relates to polynucleotides encoding a hybrid polypeptide  
 CC comprising a signal sequence and three segments that are either  
 CC contiguous or separated by a spacer amino acid or spacer peptide. The  
 CC invention specifically details polynucleotides encoding a polypeptide  
 CC peptide where the peptide segments are tumor antigens or a naturally  
 CC occurring protein of a pathogenic agent. The polypeptide peptides exhibit  
 CC antiviral and immunostimulatory activity. The polynucleotide and  
 CC polypeptide peptides are useful for eliciting an immune response in a  
 CC mammal. The polynucleotide and protein are useful as vaccines for  
 CC treating tumors and pathogenic infections. The polynucleotide is also  
 CC useful for preventing or treating human papillomavirus (HPV)-associated  
 CC diseases, particularly exophytic condyloma, flat condyloma, cervical  
 CC cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV  
 CC infection, cervical dysplasia, high grade squamous intraepithelial  
 CC lesions, and anal HPV infection. The polynucleotide and polypeptide are  
 CC useful for generating or enhancing prophylactic or therapeutic immune  
 CC response against pathogens, tumors or autoimmune diseases in a  
 CC population of individuals having diverse MHC allotypes, as positive  
 CC controls in T cell stimulation assays in vitro, and as tools to  
 CC understand processing of epitopes within cells. Peptides  
 CC AAB95894 - AAB96037 and AAB96044 - AAB96048 represent major  
 CC histocompatibility complex I (MHC I) associated tumor and pathogen  
 CC antigens. The peptides can be used as part of the polypeptide proteins of  
 CC the invention. Also included are examples of the polypeptide proteins of  
 CC represented by AAB96050 - AAB96052, and localisation signal peptides  
 CC AAB96038 - AAB96043 and AAB96049 which can be used in the construction of  
 CC the polypeptide peptides.  
 XX  
 SQ Sequence 236 AA;

Query Match 100.0%; Score 96; DB 22; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-08;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KATLODIVLHLEPQNEIPV 19  
 |||||  
 DB 187 katlgdivlhleqpneipv 205

RESULT 11  
 AAR27727  
 ID AAR27727 standard; Protein: 271 AA.

AC AAR27727;

DT 09-MAR-1993 (first entry)

DE HPV 18 E7 protein fragment.

KW Virus vector: vaccinia virus; papillomavirus; HPV; human;  
 amplification; immunotherapeutic.

OS Human papillomavirus 18.

FT key Location/Qualifiers  
 FT Peptide 164..269  
 /note="HPV-18 E7 protein"

PN W0921636-A.  
 PD 01-OCT-1992.  
 XX

PF 10-MAR-1992; 92WO-GB00424.  
 XX  
 PR 14-MAR-1991; 91GB-0005383.  
 XX  
 PA (IMMU ) IMMUNOLOGY LTD.  
 XX  
 PI Bournell MEG, Inglis SC, Munro AJ;  
 XX  
 DR WPI: 1992-349219/42.  
 DR N-PSDB; AA029390.

XX  
 PT Recombinant virus vectors encoding human papillomavirus proteins  
 PT - for treating and vaccinating against HPV infections and  
 PT conditions caused by them, such as cervical cancer  
 XX  
 PS Disclosure; Fig 1b; 83pp; English.

CC The fragment of DNA contg. the HPV-18 E6/E7 coding region was  
 CC prep'd. by PCR from plasmid pR322/HPV16 (Boshart et al., EMBO J. 3:  
 CC 1151) using oligonucleotides S01 and S02. The prod. of the  
 CC second reading frame is the HPV-18 E7 protein whereas the third  
 CC reading frame encodes HPV-18 E6. The E6 and E7 ORFs are fused  
 CC together to form a single continuous ORF via site directed mutagenesis  
 CC and the immortalising potential of E7 is removed by altering two key  
 CC codons of the HPV E7 sequence. The single ORF of HPV-18 E6/E7 may be  
 CC inserted into vaccinia virus DNA at neutral sites (pref. by inserting  
 CC two sets of the DNA in opposite orientations to overcome the problem  
 CC of intertypic recombination) to make a recombinant virus vector for  
 CC use immunotherapeutically to activate cells of the immune system  
 CC against HPV. See also AAR27723-43.  
 XX  
 SQ Sequence 271 AA;

Query Match 100.0%; Score 96; DB 13; Length 271;  
 Best Local Similarity 100.0%; Pred. No. 5e-08;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KATLODIVLHLEPQNEIPV 19  
 |||||  
 DB 168 katlgdivlhleqpneipv 186

RESULT 12  
 AAY25386  
 ID AAY25386 standard; Protein: 383 AA.

AC AAY25386;

DT 06-SEP-1999 (first entry)

DE HPV fusion protein DL/3-E6/E7-His/HPV18.

KW Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;  
 immunological fusion partner; Cpg oligonucleotide; immune response;  
 HPV antigen; prevention; treatment.

OS Synthetic.

OS Human papillomavirus.

PN W09933868-A2.

PD 08-JUL-1999.

PF 18-DEC-1998; 98WO-EP08563.

PR 24-DEC-1997; 97GB-0027262.

PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

PI Dalemans WLJ, Gerard CMG;

DR WPI: 1999-405485/34.

DB 118 katlqdivihlepneipv 136

# RESULT 8

AAV02638  
ID AAV02638 standard; Protein: 227 AA.

AC AAV02638;

DF 22-JUN-1999 (first entry)

DE Prot.D1/3-E7-His/HPV18 protein.

XX Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;

KM tumour; lesion; benign; malignant; virus; infection.

OS Chimeric - Haemophilus influenzae.

OS Chimeric - Human papillomavirus.

PN WO9910375-A2.

PD 04-MAR-1999.

PF 17-AUG-1998; 98WO-EP05285.

PR 22-AUG-1997; 97GB-0017953.

PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.

PI Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;

PI Lombardo-Bencheikh A;

DR WPI: 1999-190587/16.

DR N-PSDB: AAX29787.

XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for

PT treatment or prophylaxis of HPV induced lesions

PS Disclosure; Fig 16; 95pp; English.

XX This sequence represents a chimeric E6 or E7 protein or E6/E7 fusion

CC protein from Human papillomavirus (HPV) linked to an immunological fusion

CC partner, in this case, a fragment of the Haemophilus influenzae B

CC protein D. The sequence also contains a histidine tag at the C-terminus

CC of the encoded protein. The protein can be used in a vaccine, for

CC immuno-therapeutically treating HPV induced tumour lesions (benign or

CC malignant) and preventing HPV viral infection.

CC Sequence 227 AA;

Query Match 100.0%; Score 96; DB 20; Length 227;

Best Local Similarity 100.0%; Pred. No. 4.1e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLQDIVIHLEPNEIPV 19

DB 118 katlqdivihlepneipv 136

RESULT 9

AAV02640

ID AAV02640 standard; Protein: 227 AA.

OS Chimeric - Human papillomavirus.

OS Chimeric - Haemophilus influenzae.

PN WO9910375-A2.

PD 04-MAR-1999.

PF 17-AUG-1998; 98WO-EP05285.

PR 22-AUG-1997; 97GB-0017953.

PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.

PI Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;

PI Lombardo-Bencheikh A;

DR WPI: 1999-190587/16.

DR N-PSDB: AAX29788.

PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for

PT treatment or prophylaxis of HPV induced lesions

PS Disclosure; Fig 20; 95pp; English.

XX This sequence represents a chimeric E6 or E7 protein or E6/E7 fusion

CC protein from Human papillomavirus (HPV) linked to an immunological fusion

CC partner, in this case, a fragment of the Haemophilus influenzae B

CC protein D. The sequence also contains a histidine tag at the C-terminus

CC of the encoded protein. The protein can be used in a vaccine, for

CC immuno-therapeutically treating HPV induced tumour lesions (benign or

CC malignant) and preventing HPV viral infection.

CC Sequence 227 AA;

Query Match 100.0%; Score 96; DB 20; Length 227;

Best Local Similarity 100.0%; Pred. No. 4.1e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLQDIVIHLEPNEIPV 19

DB 118 katlqdivihlepneipv 136

RESULT 10

AAB96052

ID AAB96052 standard; Peptide: 236 AA.

AC AAB96052;

DF 25-JUN-2001 (first entry)

DE Polypeptide polypeptide SEQ ID 157.

XX Epitope; tumour antigen; antiviral; immunostimulatory; cervical cancer;

KM human papillomavirus-associated disease; condyloma; cervical dysplasia;

OS cervical dysplasia; major histocompatibility complex; MHC I.

OS Synthetic.

WO200119408-A1.

OS Chimeric - Human papillomavirus.

OS Chimeric - Haemophilus influenzae.

PN WO9910375-A2.

PD 04-MAR-1999.

PF 17-AUG-1998; 98WO-EP05285.

PR 22-AUG-1997; 97GB-0017953.

PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.

PI Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;

PI Lombardo-Bencheikh A;

DR WPI: 1999-190587/16.

DR N-PSDB: AAX29788.

PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for

PT treatment or prophylaxis of HPV induced lesions

PS Disclosure; Fig 20; 95pp; English.

XX This sequence represents a chimeric E6 or E7 protein or E6/E7 fusion

CC protein from Human papillomavirus (HPV) linked to an immunological fusion

CC partner, in this case, a fragment of the Haemophilus influenzae B

CC protein D. The sequence also contains a histidine tag at the C-terminus

CC of the encoded protein. The protein can be used in a vaccine, for

CC immuno-therapeutically treating HPV induced tumour lesions (benign or

CC malignant) and preventing HPV viral infection.

CC Sequence 227 AA;

Query Match 100.0%; Score 96; DB 20; Length 227;

Best Local Similarity 100.0%; Pred. No. 4.1e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLQDIVIHLEPNEIPV 19

DB 118 katlqdivihlepneipv 136

RESULT 10

AAB96052

ID AAB96052 standard; Peptide: 236 AA.

AC AAB96052;

DF 25-JUN-2001 (first entry)

DE Polypeptide polypeptide SEQ ID 157.

XX Epitope; tumour antigen; antiviral; immunostimulatory; cervical cancer;

KM human papillomavirus-associated disease; condyloma; cervical dysplasia;

OS cervical dysplasia; major histocompatibility complex; MHC I.

OS Synthetic.

WO200119408-A1.

CC occurring protein of a pathogenic agent. The polypeptide peptides exhibit  
CC antiviral and immunostimulatory activity. The polynucleotide and  
CC polypeptide peptides are useful for eliciting an immune response in a  
CC mammal. The polynucleotide and protein are useful as vaccines for  
CC treating tumours and pathogenic infections. The polynucleotide is also  
CC useful for preventing or treating human papillomavirus (HPV)-associated  
CC diseases, particularly exophytic condyloma, flat condyloma, cervical  
CC cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV  
CC infection, cervical dysplasia, high grade squamous intraepithelial  
CC lesions, and anal HPV infection. The polynucleotide and polypeptide are  
CC useful for generating or enhancing prophylactic or therapeutic immune  
CC response against pathogens, tumours or autoimmune diseases in a  
CC population of individuals having diverse MHC alleles, as positive  
CC controls in T cell stimulation assays in vitro, and as tools to  
CC understand processing of epitopes within cells. Peptides  
CC AAB96038 - AAB96037 and AAB96044 - AAB96048 represent major  
CC histocompatibility complex I (MHC I) associated tumour and pathogen  
CC antigens. The peptides can be used as part of the polypeptide proteins of  
CC the invention. Also included are examples of the polypeptide proteins  
CC represented by AAB96050 - AAB96052, and localisation signal peptides  
CC AAB96038 - AAB96043 and AAB96049 which can be used in the construction of  
CC the polypeptide peptides.

XX Sequence 119 AA:

Query Match 100.0%; Score 96; DB 22; Length 119;  
Best Local Similarity 100.0%; Pred. No. 2e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KATLODIVLHPONEIPV 19  
|  
Db 70 katlgdivlhpeqneipv 88

RESULT 6

AAV25382 standard; Protein: 227 AA.

AAV25382;

XX 06-SEP-1999 (first entry)

DE HPV fusion protein DI/3-E7-His/HPV18.

XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;

KM immunological fusion partner; Cpg oligonucleotide; immune response;

KW HPV antigen; prevention; treatment; ss.

XX Synthetic.

OS Human papillomavirus.

XX W09933868-A2.

XX 08-JUL-1999.

XX 18-DEC-1998; 98WO-EP08563.

XX 24-DEC-1997; 97GB-0027262.

XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Dalemans WLJ, Gerard CMG;

XX WPI: 1999-405485/34.

XX N-PSDB; AAX78798.

XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to

XX induce immune response to HPV

XX Example IX; Page 57; 62pp; English.

XX AAX78791-x78801 represent nucleic acid sequences which encode novel

CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from  
CC HPV (represented in AAV25375-Y25386). These constructs are optionally  
CC linked to an immunological fusion partner and an immunomodulatory Cpg  
CC oligonucleotide. The products of the invention can be used to induce an  
CC immune response in a patient to an HPV antigen. They can also be used  
CC for preventing or treating HPV induced tumours.

XX Sequence 227 AA:

Query Match 100.0%; Score 96; DB 20; Length 227;  
Best Local Similarity 100.0%; Pred. No. 4.1e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KATLODIVLHPONEIPV 19  
|  
Db 118 katlgdivlhpeqneipv 136

RESULT 7

AAV25384 standard; Protein: 227 AA.

AAV25384;

XX 06-SEP-1999 (first entry)

DE HPV fusion protein DI/3-E7-Mutated(C27G,E29Q)/HPV18.

XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;

KM immunological fusion partner; Cpg oligonucleotide; immune response;

KW HPV antigen; prevention; treatment.

XX Synthetic.

OS Human papillomavirus.

XX W09933868-A2.

XX 08-JUL-1999.

XX 18-DEC-1998; 98WO-EP08563.

XX 24-DEC-1997; 97GB-0027262.

XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Dalemans WLJ, Gerard CMG;

XX WPI: 1999-405485/34.

XX N-PSDB; AAX78799.

XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to

XX induce immune response to HPV

XX Example X; Page 58-59; 62pp; English.

XX AAX78791-x78801 represent nucleic acid sequences which encode novel

XX constructs comprising an E6 or E7 protein or E6/E7 fusion protein from

XX HPV (represented in AAV25375-Y25386). These constructs are optionally

XX linked to an immunological fusion partner and an immunomodulatory Cpg

XX oligonucleotide. The products of the invention can be used to induce an

XX immune response in a patient to an HPV antigen. They can also be used

XX for preventing or treating HPV induced tumours.

XX Sequence 227 AA:

Query Match 100.0%; Score 96; DB 20; Length 227;  
Best Local Similarity 100.0%; Pred. No. 4.1e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KATLODIVLHPONEIPV 19  
|  
Db 118 katlgdivlhpeqneipv 136

CC suppressor p16INK4a, inhibits kinase activity and results in growth  
CC arrest. Rb inactivators can selectively and reversibly inactivate an  
CC Rb/p16INK4a pathway, especially an Rb/p16INK4a pathway. The oncoprotein MDM2  
CC is a cellular inhibitor of Rb/p16INK4a function and the p53 tumour suppressor  
CC and can also be used in the methods. Other molecules which can be used  
CC include cdk4 or cdk6 mutants. In particular, a cdk4 mutant is one which  
CC differs from at one or more of residues R22, R24, H95 and/or D97.  
CC Additional constructs include a papilloma virus E7 protein, or other  
CC viral oncoprotein which bypasses Rb and/or p53. Antisense constructs of  
CC the Rb and p16INK4a genes may also be used. The methods are useful for  
CC increasing the proliferative capacity of cells. The cells are  
CC subsequently of use in pharmaceutical and cosmetic preparations used to  
CC treat conditions related to (premature) ageing, e.g. macular degeneration  
CC and arteriosclerosis. The cells can also be used to replace tumour cell  
CC lines in vitro and for studies on biochemical and physiological aspects  
CC of growth and differentiation. Long lived (immortal) cells could also be  
CC of use in the production of normal or genetically engineered  
CC biotechnology products.

CC XX Sequence 105 AA;

Query Match 100.0%; Score 96; DB 21; Length 105;  
Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLDIVLHLEPQNEIPV 19  
|||||  
Db 5 katldgdivlhlepqneipv 23

RESULT 4  
AAB98428  
ID AAB98428 standard; Protein: 105 AA.  
XX -

AC AAB98428;

DT 22-AUG-2001 (first entry)

DE Human papillomavirus protein HPV18 E7.

KM Human papillomavirus; human leukocyte antigen; HLA; immune response;  
KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;  
KW neoplastic growth; antiviral.

OS Human papillomavirus.

PN WO200141799-A1.

PD 14-JUN-2001.

PF 11-DEC-2000; 2000WO-US33549.

PR 10-DEC-1999; 99US-0172705.

PR 15-AUG-2000; 2000US-0641528.

PA (EPTM-) EPIMUNE INC.

PI Sette A, Sidney J, Southwood S, Chesnut R, Cells E, Grey HW;

DR WPI; 2001-381497/40.

PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for  
PT treating HPV infections -

PS Disclosure: Page 22; 756pp; English.

XX The present invention describes an isolated prepared human papillomavirus  
CC (HPV) epitope (I). (I) has antiviral activity, and can be used in  
CC vaccine production. Peptides and corresponding nucleic acid compositions  
CC from the present invention are useful for stimulating an immune response  
CC to HPV by stimulating the production of CTL or HTL responses,  
CC specifically in the treatment or prophylaxis of HPV infection, in persons

CC who have not manifested symptoms e.g. genital warts or neoplastic growth.  
CC The peptides can also be used in a tetramer staining assay to assess  
CC peripheral blood mononuclear cells for the presence of antigen-specific  
CC CTLs following exposure to a pathogen or immunogen, and as reagents to  
CC evaluate immune recall responses or evaluate the efficacy of a vaccine.  
CC The vaccine compositions are useful for removing warts or treating HPV  
CC infections. The epitopes for inclusion in an epitope-base vaccine may  
CC be selected from conserved regions of viral or tumour-associated  
CC antigens, which reduces the likelihood of escape mutants, also  
CC immunosuppressive epitopes that may be present in whole antigens can be  
CC avoided with the use of epitope-base vaccines. An additional advantage  
CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
CC the composition of the epitopes achieving enhanced immunogenicity, the  
CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
CC to AAB8477 represent polypeptide sequences used in the exemplification  
CC of the present invention.

CC XX Sequence 105 AA;

Query Match 100.0%; Score 96; DB 22; Length 105;  
Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATLDIVLHLEPQNEIPV 19  
|||||  
Db 5 katldgdivlhlepqneipv 23

RESULT 5  
AAB96051  
ID AAB96051 standard; Peptide: 119 AA.  
XX -

AC AAB96051;

DT 25-JUN-2001 (first entry)

DE Polypeptide polypeptide SEQ ID 159.

KM Epitope; tumour antigen; antiviral; immunostimulatory; cervical cancer;  
KW human papillomavirus-associated disease; condyloma; cervical dysplasia;  
KW cervical dysplasia; major histocompatibility complex; MHC I.

OS Synthetic.

PN WO200119408-A1.

PD 22-MAR-2001.

PF 18-SEP-2000; 2000WO-US25559.

PR 16-SEP-1999; 99US-0154665.

PR 16-SEP-1999; 99US-0398534.

PR 09-DEC-1999; 99US-0169846.

PR 09-DEC-1999; 99US-0458173.

PA (ZYCO-) ZYCOS INC.

PI Hedley ML, Urban RC, Chicz RM;

DR WPI; 2001-265996/27.

PT Novel nucleic acids encoding polypeptide polypeptides containing  
PT multiple epitopes from one or more proteins, useful for treating tumors  
PT and as vaccines against pathogenic agents -

PS Disclosure: Fig 4; 64pp; English.

XX This invention relates to polynucleotides encoding a hybrid polypeptide  
CC comprising a signal sequence and three segments that are either  
CC contiguous or separated by a spacer amino acid or spacer peptide. The  
CC invention specifically details polynucleotides encoding a polypeptide  
CC peptide where the peptide segments are tumour antigens or a naturally



PT multiple epitopes from one or more proteins, useful for treating tumors  
 PT and as vaccines against pathogenic agents  
 PS Claim 46; Page 10; 64pp; English.

CC This invention relates to polynucleotides encoding a hybrid polypeptide  
 CC comprising a signal sequence and three segments that are either  
 CC continuous or separated by a spacer amino acid or spacer peptide. The  
 CC invention specifically details polynucleotides encoding a polypeptide  
 CC peptide where the peptide segments are tumour antigens or a naturally  
 CC occurring protein of a pathogenic agent. The polypeptide peptides exhibit  
 CC antiviral and immunostimulatory activity. The polynucleotide and  
 CC polypeptide peptides are useful for eliciting an immune response in a  
 CC mammal. The polynucleotide and protein are useful as vaccines for  
 CC treating tumours and pathogenic infections. The polynucleotide is also  
 CC useful for preventing or treating human papillomavirus (HPV)-associated  
 CC diseases, particularly exophytic condyloma, flat condyloma, cervical  
 CC cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV  
 CC infection, cervical dysplasia, high grade squamous intraepithelial  
 CC lesions, and anal HPV infection. The polynucleotide and polypeptide are  
 CC useful for generating or enhancing prophylactic or therapeutic immune  
 CC response against pathogens, tumours or autoimmune diseases in a  
 CC population of individuals having diverse MHC allotypes, as positive  
 CC controls in T cell stimulation assays in vitro, and as tools to  
 CC understand processing of epitopes within cells. Peptides  
 CC AAB95894 - AAB96037 and AAB96044 - AAB96048 represent major  
 CC histocompatibility complex I (MHC I) associated tumour and pathogen  
 CC antigens. The peptides can be used as part of the polypeptide proteins  
 CC the invention. Also included are examples of the polypeptide proteins  
 CC represented by AAB96030 - AAB96052, and localisation signal peptides  
 CC AAB96038 - AAB96043 and AAB96049 which can be used in the construction of  
 CC the polypeptide peptides.

XX Sequence 19 AA:

Query Match 100.0%; Score 96; DB 22; Length 19;

Best Local Similarity 100.0%; Pred. NO. 2.6e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KATLQDIYVLEHPQNETPV 19  
 DB 1 KATLQDIYVLEHPQNETPV 19

RESULT 2

AAR15583  
 ID AAR15583 standard; Protein: 30 AA.

XX AAR15583;

DT 02-MAR-1992 (first entry)

DE Immunopeptide #2 derived from HPV18 E7 peptide.

KM cervical cancer; cervical intraepithelial neoplasia; CIN; wart;  
 KM squamous cell carcinoma; ELISA; HPV 18.

XX Synthetic.

PN MO9118294-A.

XX 28-NOV-1991.

PF 13-MAY-1991; 91WO-SE00335.

PR 11-MAY-1990; 90SE-0001705.

XX (MEDS-) MEDSCAND AB.

PI Dillner J, Dillner L, Cheng HM;

DR WPI: 1991-369390/50.

XX Diagnosis of human papilloma virus infection and PV-carrying  
 PT tumours - using synthetic peptide(s) to detect virus specific  
 PT antigen-antibody complexes by immunoassay

XX Disclosure; Page 39; 72pp; English.

CC This is one of a large number of peptides which have been  
 CC synthesised on the basis of the amino acid sequences for the E2, E4,  
 CC E7, L1 or L2 proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The  
 CC selection of peptide sequences was based on the assumption that an  
 CC immunoreactive region might be situated in the same relative region  
 CC of a protein from different HPV types. The peptides were used in  
 CC diagnostic immunoassays to detect HPV-infection.  
 CC See AAR15523-R15601.

XX Sequence 30 AA:

Query Match 100.0%; Score 96; DB 12; Length 30;

Best Local Similarity 100.0%; Pred. No. 4.3e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KATLQDIYVLEHPQNETPV 19  
 DB 4 KATLQDIYVLEHPQNETPV 22

RESULT 3

AAY96569  
 ID AAY96569 standard; Protein: 105 AA.

XX AAY96569;

DT 12-SEP-2000 (first entry)

DE Human papilloma virus 18 E7 polypeptide.

KM hEST2; telomerase; catalytic subunit; reverse transcriptase; life-span;  
 KM retinoblastoma; p53; tumour suppressor; inhibitor; arteriosclerosis;

KM proliferation; immortal; tumour therapy; macular degeneration; activator;  
 KM INK4; human papillomavirus; HPV-18; E7.

XX Human papillomavirus type 18.

PN WO200031238-A2.

XX 02-JUN-2000.

XX 24-NOV-1999; 99WO-US27907.

PR 25-NOV-1998; 98US-0109891.

PR 17-FEB-1999; 99US-0120549.

XX (GENE-) GENETICA INC.

PI Hannon GJ, Beach DH;

DR WPI: 2000-400055/34.

PT New method for increasing the proliferative capacity of cell lines  
 PT comprises administering agents reversibly activating telomerase  
 PT activity and reversibly inactivating Rb/INK4 and/or p53 pathways useful  
 PT in treating age related diseases

PS Claim 5; Page 12; 123pp; English.

CC The invention concerns methods and reagents for extending the life-span,  
 CC e.g. the number of mitotic divisions, of a cell. The method relies on  
 CC activation of a telomerase activity and inhibition of one or both of a  
 CC retinoblastoma (Rb)/INK4 pathway or a p53 pathway. Phosphorylation of Rb  
 CC by cyclin-dependent kinases, cdk4 and cdk6, releases the cells into the  
 CC division cycle. Binding of INK4 family members, e.g. the tumour

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 09:59:19 ; Search time 66.5 Seconds  
(without alignments)  
21.164 Million cell updates/sec

Title: US-09-664-225-154

Perfect score: 96  
Sequence: 1 KATLDIVLHLEPQNEIPV 19

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.GeneSeq\_1101.\*  
1: /SID8/gcgdata/geneSeq/geneSeq/AA1980.DAT.\*  
2: /SID8/gcgdata/geneSeq/geneSeq/AA1981.DAT.\*  
3: /SID8/gcgdata/geneSeq/geneSeq/AA1982.DAT.\*  
4: /SID8/gcgdata/geneSeq/geneSeq/AA1983.DAT.\*  
5: /SID8/gcgdata/geneSeq/geneSeq/AA1984.DAT.\*  
6: /SID8/gcgdata/geneSeq/geneSeq/AA1985.DAT.\*  
7: /SID8/gcgdata/geneSeq/geneSeq/AA1986.DAT.\*  
8: /SID8/gcgdata/geneSeq/geneSeq/AA1987.DAT.\*  
9: /SID8/gcgdata/geneSeq/geneSeq/AA1988.DAT.\*  
10: /SID8/gcgdata/geneSeq/geneSeq/AA1989.DAT.\*  
11: /SID8/gcgdata/geneSeq/geneSeq/AA1990.DAT.\*  
12: /SID8/gcgdata/geneSeq/geneSeq/AA1991.DAT.\*  
13: /SID8/gcgdata/geneSeq/geneSeq/AA1992.DAT.\*  
14: /SID8/gcgdata/geneSeq/geneSeq/AA1993.DAT.\*  
15: /SID8/gcgdata/geneSeq/geneSeq/AA1994.DAT.\*  
16: /SID8/gcgdata/geneSeq/geneSeq/AA1995.DAT.\*  
17: /SID8/gcgdata/geneSeq/geneSeq/AA1996.DAT.\*  
18: /SID8/gcgdata/geneSeq/geneSeq/AA1997.DAT.\*  
19: /SID8/gcgdata/geneSeq/geneSeq/AA1998.DAT.\*  
20: /SID8/gcgdata/geneSeq/geneSeq/AA1999.DAT.\*  
21: /SID8/gcgdata/geneSeq/geneSeq/AA2000.DAT.\*  
22: /SID8/gcgdata/geneSeq/geneSeq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	19	22	AA895965
2	96	100.0	30	12	AA815583
3	96	100.0	105	21	AAV96569
4	96	100.0	105	22	AA898428
5	96	100.0	119	22	AA896051
6	96	100.0	227	20	AAV25382
7	96	100.0	227	20	AAV25384
8	96	100.0	227	20	AAV02638
9	96	100.0	227	20	AAV02640
10	96	100.0	236	22	AA896052
11	96	100.0	271	13	AA827727

12	96	100.0	383	20	AAV25386	HPV fusion protein
13	96	100.0	383	20	AAV02642	Prot.D1/3-E6-E7-H1
14	76.5	79.7	106	22	AA898441	Human papillomavirus
15	63	65.6	48	12	AA814752	Seroreactive epitope
16	55	57.3	10	15	AA861625	Peptide fragment (
17	55	57.3	10	22	AA896027	HPV 18 E7 A1 MHC-b
18	55	57.3	105	22	AA898453	Human papillomavirus
19	48	50.0	10	22	AA896029	HPV 18 E7 A2 MHC-b
20	46	47.9	23	17	AAW06095	Human TATA-binding
21	46	47.9	23	18	AAW25013	TATA-binding prote
22	45	46.9	1513	22	AAU03531	Human protein Xina
23	44	45.8	9	14	AA843757	MHC Class I allele
24	44	45.8	9	15	AA859105	Peptide fragment (
25	44	45.8	9	18	AAW39675	HPV18 E7 peptide (
26	44	45.8	9	19	AAW54742	Peptide from HPV 1
27	44	45.8	9	20	AAV17045	HPV antigenic pept
28	44	45.8	9	21	AAV44955	Human papilloma vi
29	44	45.8	9	22	AA896028	HPV 18 E7 A2 MHC-b
30	44	45.8	20	18	AAW19962	HPV hybrid protein
31	44	45.8	20	18	AAW19963	HPV E7 region prod
32	44	45.8	99	14	AA848202	Human gene 14-enco
33	44	45.8	231	22	AA871377	Arabidopsis thalia
34	44	45.8	394	21	AA857980	Human secreted pro
35	44	45.8	613	22	AA871381	Human gene 14-enco
36	44	45.8	613	22	AA893097	Human protein sequ
37	44	45.8	613	22	AA894805	Human protein sequ
38	44	45.8	699	22	AA894300	Human protein sequ
39	44	45.8	714	21	AA832994	Arabidopsis thalia
40	44	45.8	92	21	AA802359	Human secreted pro
41	43	44.8	206	22	AA893859	Human protein sequ
42	42.5	44.3	400	22	AAE05359	Mouse insulin-like
43	42	43.8	65	22	AAW21501	Peptide #7935 enco
44	42	43.8	65	22	AAW37756	Peptide #11793 enc
45	42	43.8	98	22	AA898396	Human papillomavir

#### ALIGNMENTS

RESULT 1	
ID	AA895965 standard; Peptide: 19 AA.
AC	AA895965;
DT	25-JUN-2001 (first entry)
DE	HPV 18 E6 protein fragment SEQ ID 154.
KW	Epitope; tumour antigen; antiviral; immunostimulatory; cervical cancer; human papillomavirus-associated disease; condyloma; cervical dysplasia; cervical dysplasia; major histocompatibility complex; MHC I.
KW	Human papillomavirus.
OS	Human papillomavirus.
PN	WO200119408-A1.
PD	22-MAR-2001.
PF	18-SEP-2000; 2000WO-US25559.
PR	16-SEP-1999; 99US-0154665.
PR	16-SEP-1999; 99US-0398534.
PR	09-DEC-1999; 99US-0169846.
PR	09-DEC-1999; 99US-0458173.
PA	(ZYCO-) ZYCOS INC.
PI	Hedley ML, Urban RC, Chicz RM;
XX	WPI, 2001-265996/27.
XX	Novel nucleic acids encoding polypeptide polypeptides containing